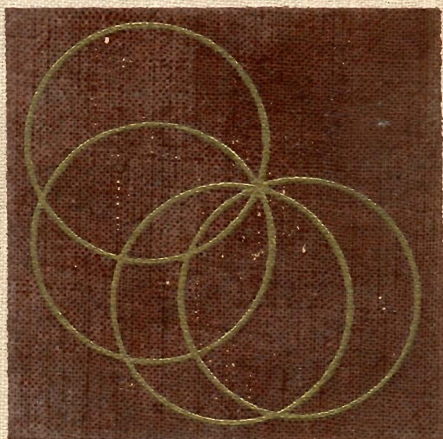


И.И.РѢВЗИН

М Е Т О Д  
М О Д Е Л И Р О В А Н И Я



И Т И П О Л О Г И Я  
С Л А В Я Н С К И Х  
Я З Ы К О В

АКАДЕМИЯ НАУК СССР  
ИНСТИТУТ СЛАВЯНОВЕДЕНИЯ

*И. И. Ревзин*

МЕТОД  
МОДЕЛИРОВАНИЯ  
И ТИПОЛОГИЯ  
СЛАВЯНСКИХ ЯЗЫКОВ



*Издательство «Наука»*

МОСКВА 1967

Монография представляет собой значительно переработанный и дополненный новым материалом из славянской типологии вариант книги того же автора «Модели языка», выпущенной Издательством АН СССР в 1962 г.

В книге излагаются новейшие результаты в области лингвистической теории моделей, в частности теории порождающих грамматик.

Работа содержит богатый фактический языковой материал, что делает ее интересной для широких кругов специалистов.

Ответственный редактор

ВЯЧ. ВС. ИВАНОВ

В книге «Модели языка» автор изложил свою концепцию лингвистического моделирования, пытаясь обобщить с определенной точки зрения работу, проделанную в этой области советскими учеными.

После выхода в свет книги «Модели языка» автор получил многочисленные отзывы и замечания от отдельных лиц и коллективов (наиболее важными для автора были замечания А. В. Гладкого и А. А. Зализняка, а также замечания Ю. Сирка). Кроме того, в печати появился ряд рецензий и обзоров<sup>1</sup>, посвященных этой книге, а также упоминания об отдельных проблемах, разбираемых в ней.

С другой стороны, у нас за истекшие годы развивались как исследования по моделированию морфологии (для автора были наиболее важны работы А. А. Зализняка и О. Г. Карпинской), так и в особенности исследования по моделированию синтаксиса (здесь автор опирался в первую очередь на новые исследования М. И. Белецкого, А. В. Гладкого, И. А. Мельчука, Е. В. Падучевой, С. Я. Фитиллова, Ю. А. Шрейдера).

Из зарубежных работ для автора были важны те работы, которые примыкают к так называемому «теоретико-множественному» моделированию языка<sup>2</sup>. Здесь следует назвать многочислен-

<sup>1</sup> И. М. Яглом [рец.] И. И. Ревзин. Модели языка. М., 1962. — «Вестник АН СССР», № 7. М., 1963, стр. 128—129; L. Antal — «Linguistics», 1964, № 3; K. Berka, P. Novák. Východ fonologických a gramatických pojmů pomocí pojmů teorie množin (nad knihou I. I. Revzina «Modeli jazyka». М., 1962). «Slovo a slovesnost», 1963, № 2, стр. 133—139; L. Zgusta — «Archiv orientální», № 31, 1963, стр. 687—688; S. Marcus — «Analele romîno-sovietice. Automatică și cibernetică», 3, 1963, стр. 196—199; F. Papp — «Acta Linguistica Hungarica», XV, 1965, стр. 181.

<sup>2</sup> Термин «теоретико-множественное моделирование» имеет скорее исторический характер, поскольку восходит к первой у нас работе в этой области, а именно работе О. С. Кулагиной. Позднее (например, в работах Д. Н. Ленского и М. Новотного) было выяснено, что подход Кулагиной является по существу алгебраическим [может быть, вообще термин «алгебраическая лингвистика», применяемый чешскими лингвистами для

ные работы румынского математика С. Маркуса, работы ряда чешских ученых (Фр. Данеш, М. Новотный, П. Новак, П. Сгалл), серию исследований по проективности языка, а также последние по порядку, но не по важности работы, вызванные исследованиями Н. Хомского. Особенно большое впечатление на автора произвело обсуждение идей этого ученого на американском симпозиуме по прикладной математике, состоявшемся в апреле 1960 г. в Нью-Йорке<sup>3</sup>.

Поскольку в период написания книги «Модели языка» автор еще не придавал идеям Хомского должного значения, то в рецензии на указанный сборник трудов симпозиума он считал своим долгом специально остановиться на месте этих идей в современном языкознании<sup>4</sup>.

Большое значение для развития общего идейного аппарата теории лингвистических моделей имели работы по семиотике, проводившиеся сектором структурной типологии славянских языков и отчасти нашедшие свое отражение в тезисах Симпозиума по структурному изучению знаковых систем (Москва, 1962). В первую очередь здесь необходимо отметить работы Вяч. Вс. Иванова и В. Н. Топорова. Автору были весьма полезны многочисленные беседы с С. К. Шаумяном об общих методологических проблемах лингвистики.

Чрезвычайно важные указания были сделаны автору Р. И. Аванесовым, Л. А. Калужниным и П. С. Кузнецовым.

Автор, таким образом, должен был учитывать как непосредственную критику своей работы, так и ту косвенную критику, которая содержалась в новых теориях и уточнениях. И та и другая критика была очень острой и именно поэтому весьма полезной для автора.

Наиболее существенные изъяны оказались в пятой главе книги «Модели языка», которая, как теперь ясно осознает автор, была совершенно неудовлетворительной (следует заметить, что

всей области неколичественного моделирования, был бы более удачным (см.: P. Sgall a kolektiv. Cesty moderní jazykovědy. Praha, 1964, стр. 74)].

<sup>3</sup> См.: «Proceedings of symposia in applied mathematics». Vol. XII. «Structure of language and its mathematical aspects», ed. R. Jakobson. Providence, 1961. Наиболее важные материалы этого симпозиума опубликованы в русском переводе (см.: «Новое в лингвистике», вып. IV. М., 1965).

<sup>4</sup> Рецензия («Word», 1963, № 3) вызвала негодование Р. А. Будагова (Р. А. Б у д а г о в. Проблемы развития языка. М., 1965, стр. 23), особенно место, где значение идей Хомского для лингвистики 60-х годов сравнивается со значением идей Соссюра для лингвистики 20-х годов. Р. А. Будагову понадобилось 50 лет, чтобы оценить место Соссюра в истории науки; через такой же промежуток времени можно было бы вернуться к оценке им Хомского. Это же место в рецензии весьма иронически цитируется В. А. Звегинцевым («один из советских языковедов даже объявил *urbi et orbi*. . .» см.: «Новое в лингвистике», вып. IV. М., 1965, стр. 393). Автор так и не понял, становится ли тон В. А. Звегинцева более снисходительным, когда он пишет о «Моделях языка» (там же, стр. 9—14).



в момент окончания книги автор был менее всего удовлетворен именно этой главой<sup>5</sup>, но он, разумеется, не отдавал себе отчета о ее действительных недостатках). Благодаря А. В. Гладкому автор получил возможность подробно разобраться в ошибках, связанных с доказательствами приводимых в пятой главе теорем. Некоторые из них оказались просто неверными (например, теоремы 5.5, 5.8, 5.9, 5.10), другие неверны, если не уточнить определения, на которых они основаны. Разумеется, можно было бы сослаться на то, что при создании некоторой новой теории, в особенности такой, где важна в первую очередь выработка системы определений, теоремы имеют смысл лишь как способ проверки операционности определений, и с этой точки зрения тот факт, что, оставаясь внутри принятой системы определений, можно доказать ложность некоторых утверждений, несомненно, имеет положительное значение (ср. замечание Бэкона о том, что «истина может возникнуть скорее из ошибки, нежели из смешения понятий»<sup>6</sup>). Во всяком случае после того, как ошибки обнаружены, их было бы нетрудно исправить. Гораздо более важной для автора была косвенная критика, состоявшая в том, что идеи пятой главы почти не нашли отклика в работах других исследователей<sup>7</sup>.

Ценность своих идей автор может проверить лишь одним способом: тем откликом, который они нашли в науке. Если идеи автора в области фонологии и морфологии прямо или косвенно разрабатываются в целом ряде работ (а в других работах автор встречает близкие по духу построения), то синтаксические концепции автора, как они были изложены в пятой главе книги, остались в стороне от магистральной линии развития теории моделей.

Именно это последнее обстоятельство заставило автора полностью пересмотреть свою синтаксическую концепцию. Основное его устремление состояло в том, чтобы увязать идеи других исследователей с его собственными взглядами на роль грамматических категорий в языке и показать важность простых и обычно отбрасываемых моделей для уточнения лингвистических проблем.

Написанная им совершенно новая глава о синтаксических моделях в гораздо большей мере, чем раньше, учитывает достигнутое в данной области, хотя и содержит весьма существенную, с точки зрения автора, модификацию обычной концепции. Автор считает наиболее важными следующие два соображения: во-первых, до сих пор идеи синтаксического моделирования в известной мере оторваны от рассмотрения лингвистической реальности,

<sup>5</sup> См.: И. И. Ревзин. Модели языка, стр. 5.

<sup>6</sup> Цит. по кн.: Э. Косериу. Синхрония, диахрония и история. Перев. с исп. «Новое в лингвистике», вып. III. М., 1963, стр. 155.

<sup>7</sup> За исключением, может быть, работы: А. В. Гладкий. Об одном способе формализации понятия синтаксической связи. «Проблемы кибернетики», вып. 11. М., 1964.

которая может быть поставлена в соответствие каждому типу модели в известной иерархии Хомского; во-вторых, автор выделил в целом ряде описываемых моделей черты, заставляющие, как он считает, излагать синтаксис не на уровне отдельных слов, как это принято в работах по моделированию синтаксиса, а на уровне синтаксических групп, как это принято в традиционной синтаксической теории.

Таким образом, и для синтаксиса остается верной основная идея других глав: идея о необходимости преемственной связи с лингвистической традицией.

Работая над этой главой, автор предполагал, что он может ограничиться ее переработкой, а в остальных главах учесть лишь отдельные замечания, сделанные ему. Таким образом, в начале речь шла о подготовке второго издания книги «Модели языка».

Однако по мере работы над другими главами для второго издания книги «Модели языка» автор все более ясно осознавал, что как сам материал, так и его трактовка все более выходят за рамки даже значительной переработки книги и что по существу написана новая книга, связанная с «Моделями языка» лишь некоторыми исходными понятиями.

Так, идеи двух параграфов «Моделей языка» (§ 28—29) выросли в две главы: о частях речи и о теоретико-множественной типологии, а материалы следующего параграфа (§ 30) в главу о моделировании грамматических категорий.

Очень сильно переработаны и остальные главы, причем после того, как стало ясно, что возникает новая книга, автор везде старался избежать повторения того, что сказано в «Моделях языка», и там, где соответствующие параграфы указанной книги представлялись автору корректно изложенными, он ограничился ссылками (кроме того, в «Приложении» дан список важнейших исправлений к книге «Модели языка»).

Почти полностью изменена первая глава, замененная двумя отдельными главами. Это изменение в первую очередь связано с недавно ставшими известными автору статьями профессора А. Ф. Лосева «О методах изложения математической лингвистики для лингвистов» и «Трудности построения теории грамматических моделей методами математической лингвистики»<sup>8</sup>. Эти две статьи общим объемом в 75 страниц машинописного текста целиком посвящены критике книги «Модели языка» и, если отвлечься от обвинений в идеализме<sup>9</sup>, были весьма полезны автору. Они

---

<sup>8</sup> См. сокращенный вариант этих статей: А. Ф. Лосев. О трудностях изложения математической лингвистики для лингвистов. — ВЯ, 1965, № 5.

<sup>9</sup> Здесь профессор Лосев смыкается с Т. А. Дегтяревой, которая не увидела в книге «Модели языка» ничего, кроме проповеди воинствующего «идеализма» (Т. А. Дегтярева. Пути развития советского языкознания, т. III. М., 1964).

показали, что в ряде мест читатель-филолог склонен принимать за определения рассуждения, направленные на установление связи понятий теории моделей с традиционно-лингвистическими.

Поскольку в книге «Модели языка» определения специально не выделялись (многие недоумения профессора Лосева связаны именно с определениями), а определения составляют суть работы, автор выделил и перенумеровал все определения (они снабжаются буквой *O*, номером главы и после точки следует номер определения в данной главе; этот же принцип принят для теорем — *T*).

В свою очередь определения разбиты на две группы: а) определения, основанные на четко определенных исходных понятиях, б) определения, не основанные на таких понятиях и не являющиеся поэтому определениями в строгом смысле слова, но важные как предварительная попытка уточнения некоторых понятий (эти определения будут даваться под звездочкой, например, *O\* 1. 1*).

Весь остальной текст можно рассматривать лишь как комментарий к приводимым определениям (как собственно определениям, так и к определениям под звездочкой).

Чтобы не возникало недоразумений по поводу отношения автора к семантике (столь характерных для А. Ф. Лосева и многих других критиков теории моделей) в книге добавлена специальная VIII глава, излагающая точку зрения автора.

Реакция профессора Лосева также показала автору, что в книге «Модели языка» были сделаны слишком сильные допущения об осведомленности филологов, интересующихся проблемами моделирования, в области логической терминологии. Хотя данная монография так же, как и «Модели языка», не является учебным пособием по «математической» лингвистике, автор счел возможным написать специальную главу, в которой вводятся такие основные логические и математические понятия, как множество, разбиение, класс, отношение, отображение, граф, алгоритм. Автор надеется, что теперь для чтения книги филологу не потребуются никакой предварительной подготовки. Отдельные лингвистические примеры, приводимые в той же главе, могут представлять интерес и для читателей, знакомых с вводимыми понятиями.

Вообще же в данной книге еще более последовательно, чем в «Моделях языка», проводится тенденция сократить логико-математический аппарат до минимума, оставив лишь наиболее общие и действительно необходимые лингвисту понятия.

Предпринята также попытка упростить терминологию и сделать ее по возможности согласованной с терминологией, употребляемой другими авторами. Автор отказался поэтому от ряда терминов, использовавшихся им ранее, и надеется, что добрая воля, проявленная им, побудит и других авторов со своей стороны сделать какие-то уступки с целью унификации терминологии (большое влияние на автора в этом отношении оказал А. В. Гладкий).



Наконец, чтобы предупредить возможные (как показали статьи А. Ф. Лосева) недоразумения, необходимо сделать одно замечание относительно способа изложения материала. Во всей книге, за исключением первой главы, каждое понятие используется лишь после того, как оно получило определение в предшествующем тексте (или введено как неопределяемое). Исключение составляет первая, вводная глава, где используется ряд понятий, определяемых лишь в дальнейшем, но уже известных лингвистам в их общем, хотя и не всегда строго определенном значении. Автор по своему опыту знает, что лингвисту легче начинать с такого изложения, ибо если бы изложение начиналось со второй главы, то оно могло бы отпугнуть многих читателей, которых автору не хотелось бы терять. Что же касается читателей, которые предъявляют повышенные претензии к формальной строгости изложения, то они могут начинать чтение со второй главы и лишь после прочтения всей книги вернуться (если такая необходимость возникнет) к первой главе.

В заключение автор хотел бы поблагодарить Вяч. Вс. Иванова, И. А. Мельчука, Т. М. Николаеву, О. Г. Ревзину (Карпинскую) и С. К. Шаумяна, ознакомившихся на разных этапах подготовки книги с отдельными разделами рукописи настоящей работы и сделавшими ряд критических замечаний, и в особенности А. В. Гладкого, согласившегося просмотреть два последовательных варианта рукописи, второй из которых значительно отличался от первого именно благодаря чрезвычайно важным замечаниям новосибирского математика. Автор глубоко обязан указанным лицам, а также всем тем, кто написал или устно сообщил ему о недостатках «Моделей языка» и тем самым содействовал написанию настоящей книги.

## ПРИНЦИПЫ ЛИНГВИСТИЧЕСКОГО МОДЕЛИРОВАНИЯ

### § 1. НЕКОТОРЫЕ СЕМИОТИЧЕСКИЕ ПОНЯТИЯ И ОБЩИЕ ПРИНЦИПЫ АНАЛИЗА ЗНАКОВЫХ СИСТЕМ

Абстрактный подход к исследованию естественных языков, связанный с идеями моделирования, привел к тому, что естественные языки начали изучать как частный случай знаковой системы, а лингвистика постепенно становится составной частью семиотики.

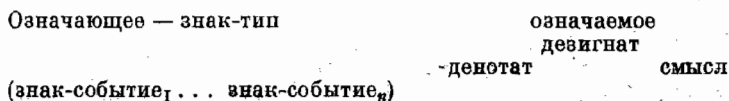
*Семиотика есть наука об общих свойствах знаковых систем и законах их функционирования вне зависимости от конкретного воплощения соответствующих знаков и области их применения.* В языкознании возникновение семиотического подхода к языку связано с именем Соссюра, с проведенным им разделением внешней лингвистики (исследования связи истории языка и истории его носителей, а также социальных условий коллектива, в котором функционирует язык) и внутренней лингвистики (исследования законов, управляющих соотношением единиц внутри языка). Соссюр увидел, что в рамках внутренней лингвистики естественный язык можно сравнить с другими совокупностями знаков, как, например, с азбукой для глухонемых, с символическими обрядами, с формами учтивости, с военными сигналами и т. д. Науку, изучающую «жизнь знаков внутри общества», Соссюр предложил назвать семиологией (позднее, однако, утвердился термин «семиотика»).

Развитие наук и усовершенствование их языков, возникновение кибернетики и необходимость создания информационных языков и языков программирования для машин, способных накапливать и перерабатывать информацию, необыкновенно расширили круг знаковых систем (языков в широком смысле слова), используемых в настоящее время. Все более ощутимой стала потребность в науке семиотике, которая дала бы возможность изучать все эти явления с единой точки зрения.

Основное семиотическое понятие — это понятие знака. Знаками являются, например, слова и морфы (минимальные значимые части слов) естественного языка, символы химического языка (например, O — атом кислорода, H — атом водорода, 3O<sub>2</sub> — три молекулы кислорода и т. д.), знаки регуляции

уличного движения и т. п. Вообще под знаком следует понимать отношение между совокупностью чувственно воспринимаемых объектов, называемой означающим и некоторой другой сущностью, называемой означаемым. Подчеркнем, что означающим как правило является не один чувственно воспринимаемый объект, например, данное начертание какого-то слова, а именно множество таких объектов (начертание разными шрифтами с использованием разного материала или даже написанное в разные моменты времени). Для различия означающего как класса и отдельной манифестации знака иногда говорят о знаке-типе (в английском sign-type) и знаке-событии (в английском sign-event). Идентификация различных знаков-событий как манифестаций одного означающего одна из важных проблем семиотики, однако в данной книге мы будем исходить из того, что отдельные знаки-события, если они не являются элементами одной знаковой последовательности (ср. § 16), отождествлены, т. е. что нам даны уже знаки-типы, иначе говоря, мы будем пользоваться допустимой в семиотике абстракцией отождествления. Что касается означаемого, то структура его, как выяснено в семиотике, еще более сложна. Здесь следует различать, с одной стороны, то, что обозначается данным означающим, или, как говорят, его значение («дезигнат»). Если дезигнат есть объект или явление описываемой предметной области, то его называют денотатом (имеются дезигнаты, не являющиеся денотатами: например, значение слова *русалка* в естественном языке или значение термина *система* в языке лингвистики). С другой стороны, принято выделять в означающем еще и смысл, т. е. способ представления дезигната в знаке. Часто дезигнат и смысл совпадают, например для морфемы *больш-*, но они могут различаться (см. ниже *Богданович* и *Социниель Д-ки*).

Структура знака может быть представлена в виде следующей схемы:



Понятия денотата и дезигната предложил различать Моррис<sup>1</sup>, значения и смысла — Фреге<sup>2</sup>. Хотя понятие смысла как способа представления объекта в знаке достаточно хорошо разработано

<sup>1</sup> Ch. Morris. Foundations of the Theory of Signs. Chicago, 1938, стр. 6—9.

<sup>2</sup> Близко к этому противопоставлению разграничение понятий «станция плана содержания», т. е. обозначаемые объекты, и «форма плана содержания», т. е. представление объекта в знаке в концепции Ельмслева (см.: Л. Ельмслев, Прологомены к теории языка. Перев. с англ. «Новое в лингвистике», вып. I. М., 1960, стр. 314 и след.).

в логике, имеются лингвистические соображения, говорящие о необходимости его уточнения. Рассмотрим следующие три примера.

1) Знаки *Богданович* и *Сочинитель Д-ки* обозначают один и тот же объект. Этот же объект можно обозначить и так: *русский поэт, который после успеха своего произведения стал подписываться как сочинитель этого произведения*, или *поэт, имевший в XVIII в. большой успех как автор поэмы «Душенька», или русский поэт, родившийся в 1743 г. и умерший в 1802*, и т. п. Как видно из примеров, этот список разных способов обозначения одного объекта ничем не замкнут и его можно было бы продолжать. Заметим, что все эти описания не зависят от русского языка и с таким же успехом могли бы быть даны, например, на французском языке.

2) Французское слово *quatre vingt* называет тот же объект, что и русское слово *восемьдесят*, но первый знак представляет этот объект как  $4 \times 20$ , а второй как  $8 \times 10$ .

3) Французское слово *assouir* обозначает то же действие, что и русские слова *прибегать* и *прибежать*. Но во французском языке это действие не представляется одним из двух способов: как ограниченное во времени или как неограниченное, в то время как в русском языке нельзя обозначить это действие, минуя оба эти способа.

Эти три примера показывают, что (по крайней мере, в естественных языках) мы имеем дело с двумя разными явлениями, которые подходят под данное выше понимание термина «смысл» как способа представления объекта в знаке. В одном случае (пример 1) способ представления зависит не от языка, а от пользующегося им, причем число таких способов, как мы видели, практически неограниченно, в остальных двух случаях способ представления объекта в знаке диктуется языком таким образом, что может быть выбран лишь один из строго ограниченного самим языком набора способов (в предельном случае, как в примере 2, такого выбора вообще уже нет, так как имеется всего один способ). В дальнейшем мы будем употреблять термин «смысл» лишь для первого случая (способ представления объекта в знаке, не зависящий от языка), а во втором случае говорить о категоризации<sup>3</sup>. Такое разграничение существенно и с точки зрения перевода,

<sup>3</sup> Понятие «категоризации» вводится как своего рода обобщение понятия «внутренней формы». Это понятие, введенное Гумбольдтом, получило глубокого семиотическое истолкование в трудах Г. Г. Шпета [см.: Г. Шпета. Внутренняя форма слова (этюды и вариации на темы Гумбольдта). М., 1927]. Ср. также толкование этого термина как сопредставление у А. Марти (см.: O. F u n k e. Innere Sprachform. Eine Einführung in A. Marty's Sprachphilosophie. Reichenberg, 1924, стр. 36). Частным случаем внутренней формы Марти считал и отнесение к определенным грамматическим категориям (O. F u n k e. Указ. соч., стр. 55—58).

а именно смысл есть то, что сохраняется при переводе<sup>4</sup>, в то время как категоризация есть способ представления объекта в знаке, вообще говоря, не сохраняющийся при переводе<sup>5</sup>.

Семиотика изучает и такие компоненты знака, которые сами не являются знаками, т. е. в которых нет отношения между означающим и означаемым, но которые служат для построения знака (эти компоненты иногда называют *фигурами*). Таковы, например, звуки речи, которые сами по себе не обязательно несут значение, но зато служат для построения определенных означающих, а именно морфов. Для компонентов знака также существенно деление на компоненты-события и компоненты-типы. Заметим, что в будущем, говоря о звуках речи, мы всюду будем иметь в виду звуки-типы.

При описании некоторой знаковой системы необходима новая знаковая система, которую принято называть *метаязыком*, а описываемая знаковая система в этом случае называется *языком-объектом*, знаки и отношения знаков которого представляют предметную область метаязыка.

Особенность семиотики состоит в том, что в ней изучаются не просто наборы знаков, а знаковые системы, т. е. такие совокупности, в которых из небольшого набора первичных знаков конструируется бесконечное число новых, сложных знаков, если деизгнаты относятся к «открытой» предметной области (таковы естественные языки) или, во всяком случае, достаточно большое число знаков, если деизгнаты относятся к «замкнутой» предметной области. Отсюда вытекает свойство иерархичности знаковых систем, проявляющееся в том, что в них имеется несколько уровней.

Между отдельными элементами сложного знака, например, между словами предложения естественного языка или символами формулы сложного химического соединения, возникают определенные отношения, которые называются *синтагматическими* и отношениями (отношения порядка, подчинения и т. п.). Иначе говоря, синтагматические отношения основаны на линейной упорядоченности отдельных знаков в тексте и на тех отношениях, которые возникают в линейно-упорядоченных цепочках, например фразах. Значение порядка элементов иллюстрируется, например, немецкими сложными словами *Planreise* (плановое путешествие) и *Reiseplan* (план путешествия), в арифметической символике 520 обозначает другой объект, чем 205.

---

<sup>4</sup> Именно так определен «смысл предложения» Черчем (А. Черч. Введение в математическую логику. Перев. с англ., т. I. М., 1960, стр. 31—32).

<sup>5</sup> Нечеткое различие смысла и категоризации в § 14 книги (И. И. Резнищ, В. Ю. Розенцвейг. Основы общего и машинного перевода. М., 1965, стр. 65) не дало возможности ясно вскрыть суть дела в соответствующем параграфе.

Кроме синтагматических отношений, в каждой знаковой системе (именно в силу того, что небольшое число знаков низшего уровня дает возможность образования очень большого числа производных знаков высшего уровня) могут изучаться так называемые парадигматические отношения, т. е. отношения, основанные на сходстве означающих и означаемых двух знаков, не находящихся в синтагматическом отношении. Частным случаем парадигматических отношений будет тождество означающего и означаемого у двух различных знаков-событий, т. е. принадлежность их к одному знаку-типу. Поскольку мы договорились заранее отождествлять такие знаки, то нас будет интересовать лишь тот случай парадигматических отношений, когда они основаны на тождестве означающих при различии означаемых — таковы явления *о м о н и м и и*, — или на тождестве означаемых, при различии означающих — таковы явления *с и н о н и м и и*, или наконец на частичном сходстве означающих (или соответственно означаемых) при одновременном частичном различии их — таковы, например, в естественных языках отношения, основанные на одинаковости окончания при разной основе или наоборот на одинаковости основы при разном окончании (в последнем случае мы имеем дело с парадигмами склонения или спряжения, что и дало повод для названия этих отношений парадигматическими). Однако парадигматические отношения не ограничиваются естественными языками. Хорошим примером парадигматических отношений могут служить отношения между знаками элементов в химическом языке, описываемые периодической системой Менделеева.

В структурной лингвистике часто обсуждался вопрос, с чего начинать описание знаковой системы: с ее парадигматики или ее синтагматики. Выяснилось, что ответ на этот вопрос зависит как от цели исследования, так и от степени овладения данным языком. С описания парадигматики можно начинать лишь в том случае, если соответствующий язык хорошо знаком исследователю. Но и в этом случае начинать с выяснения некоторых синтагматических отношений в известном смысле легче.

Парадигматические отношения (важные в том смысле, что они дают наиболее компактное представление о знаковой системе) могут изучаться на их основе, а именно определенное парадигматическое отношение можно получить и как свойство класса знаков, вступающего в одинаковые синтагматические отношения. Так, эквивалентность слов *стол* и *стул*, характеризующая определенные парадигматические отношения между ними, может быть установлена на том основании, что они употребляются во фразах в одинаковом окружении. Парадигматические отношения между символами *As* и *P* в химическом языке могут быть, в частности, установлены на том основании, что формулы для солей фосфорной и мышьяковой кислот, содержащих одинаковое число молекул



кристаллизационной воды и имеющих тождественную кристаллическую структуру, различаются только этими символами (ср.:  $\text{Na}_2\text{HAsO}_4 \cdot 12\text{H}_2\text{O}$  и  $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$ ). Исчерпывающее описание знаковой системы предполагает описание всех синтагматических и парадигматических отношений, имеющих в этой системе.

В современной структурной лингвистике отчетливо выявились два диаметрально противоположных (но, как мы увидим ниже, дополняющих друг друга) пути описания знаковой системы. Один из них, который можно называть **д е с к р и п т и в н ы м**, или **а н а л и т и ч е с к и м**, состоит в том, что при описании мы исходим из заданной нам совокупности текстов, т. е. знаковых последовательностей и области, к которой относятся деизгнаты знаков, и стремимся получить набор синтагматических и парадигматических отношений в системе. Второй путь, который можно назвать **к о н с т р у к т и в н ы м**, или **с и н т е т и ч е с к и м**, состоит в том, что мы исходим из некоторого набора синтагматических и парадигматических отношений, которые интерпретируются как правила порождения всей совокупности текстов данного языка. Каждый из этих путей плодотворен для решения задач, стоящих перед семиотикой. Так, первый путь, наиболее последовательно разработанный в так называемой дескриптивной лингвистике, привел к созданию ряда формальных процедур, применимых в первую очередь для изучения малоизученных языков, а также для дешифровки, но одновременно сыгравших большую роль и в исследовании языков, которые изучены достаточно хорошо, ибо здесь соответствующие процедуры помогают прояснить лингвистическую интуицию исследователя<sup>6</sup>, т. е. определить, какими чисто формальными критериями могут быть подтверждены факты, установленные на основе содержательных соображений.

Второй путь исследований, который оформился в структурной лингвистике, с одной стороны, на основе переосмысления некоторых понятий и методов дескриптивной лингвистики, а с другой стороны, под непосредственным влиянием математической логики, позволяет исследовать не только синтаксическую функцию знаков, но и их семантическую функцию. Это теория так называемых порождающих грамматик, т. е. устройств, содержащих конечное число правил и могущих вывести любую фразу данного языка и тем самым задать все фразы данного языка. Теория порождающих грамматик имеет важное значение для решения целого ряда проблем, носящих общесемиотический характер.

---

<sup>6</sup> О важности этой задачи для современной лингвистики см.: Вяч. Вс. Иванов. Некоторые проблемы современной лингвистики. «Народы Азии и Африки», 1963, № 4.

## § 2. О РОЛИ ЛИНГВИСТИЧЕСКИХ КАТЕГОРИЙ

Как мы уже отметили в § 1, исчерпывающее описание знаковой системы предполагает перечисление всех имеющих в ней синтагматических и парадигматических отношений. Как же отвечает этим требованиям порождающая грамматика? Синтагматические отношения, как мы увидим в главе VII, отражаются в форме правил, по которым выводится предложение. Что касается парадигматических отношений в порождающей грамматике, то явно они здесь не выражены, но косвенно они отражены в выборе метаобозначений; каждое из них представляет определенный набор лингвистических категорий, а последние и раскрывают парадигматику соответствующих слов. В этом пункте Хомский довольно резко порвал с традицией дескриптивной лингвистики, из которой он исходил, поскольку для дескриптивной лингвистики, ограничивавшей сферу языковедческого исследования пределами эмпирически изучаемых языковых фактов, единственной объективной данностью является совокупность речевых высказываний.

Что нового с семиотической точки зрения дает введение лингвистических категорий в порождающую грамматику? Знаки для категорий в грамматике как метаязыке могут по-разному соотноситься со знаками языка-объекта.

а) Знаку каждой категории соответствует отдельный знак языка-объекта — таково положение в так называемых агглютинативных языках, например казахском, где означаемому «дательный падеж, множественное число» соответствует последовательность определенных морфем (ср. *стол-дар-ға* 'столам', где *дар* обозначает множественное число а *ға* — дательный падеж). Заметим, что аналогично построены и многие слова русского языка, например глаголы в повелительном наклонении (ср. *иди-те-ка*). В этом случае соответствующие категории легко выделяются и в дескриптивной модели, и введение их в порождающую грамматику представляется вполне естественным.

б) Знаку категории не обязательно соответствует некоторый знак языка-объекта — таково положение в преимущественно флективных языках, например, для большинства слов русского языка; так здесь нет отдельного знака для означаемого «дательный падеж» и отдельного знака для означаемого «множественное число». Поэтому соответствующие категории и не могут быть непосредственно выведены в дескриптивной модели. Различие между агглютинативным и флективным типом может быть истолковано в семиотических терминах как наличие или отсутствие взаимно однозначного соответствия между знаками языка-объекта и знаками метаязыка.

При построении порождающей грамматики для языков типа русского такие именные категории, как «мужской род», «датель-

ный падеж», «одушевленность» и т. п., не говоря уже о таких абстракциях, как «род», «падеж» и т. п., не даны непосредственно в формах языка и не выводятся из них путем генерализации, а постулируются, вносятся извне, и единственное оправдание такого внесения состоит в том, что с их помощью более адекватно описываются факты языка. В результате этого в метаязыке появляются знаки, десигнаты которых представляют собой не объекты исходного языка и не непосредственно наблюдаемые отношения между ними. Специфический характер этих абстрактных объектов состоит в том, что их определение должно даваться через те отношения, в которые эти абстрактные объекты вступают с наблюдаемыми лингвистическими фактами (подробнее этот вопрос рассматривается в главе VI).

Необходимость введения категорий в порождающую модель (вне зависимости от того, выражена ли каждая из них отдельным знаком) отражает реальность категорий в языковом мышлении. Можно считать, что каждый естественный язык через свои лингвистические категории по-своему «моделирует» действительность.

Факты разной категоризации действительности настолько очевидны, что только теоретическая предвзятость заставляет некоторых критиков структурализма отвергать их. Эти лингвисты не учитывают, что отражение действительности представляет собой процесс, связанный с активной переработкой информации и ее внутренним упорядочением через систему лингвистических категорий<sup>7</sup>. Категория языка, не меняя значения (предметной соотнесенности сложного знака), меняет способ представления соответствующего объекта в знаке.

### § 3. О СООТНОШЕНИИ ТРАДИЦИОННОЙ ЛИНГВИСТИКИ И ТЕОРИИ МОДЕЛЕЙ

Поскольку многие критики современной теории лингвистических моделей и — шире — современной структурной лингвистики ставят этому направлению в упрек разрыв с традиционной лингвистикой, то автор с самого начала хотел бы уточнить свою позицию в этом вопросе.

| Попытки противопоставить «традиционную» и «структурную» лингвистику как две разные науки делались неоднократно. Эти противопоставления идут в основном по следующим направлениям: а) противопоставления по объекту исследования; б) противопоставления по цели исследования; в) противопоставления по отношению к предшествующей

---

<sup>7</sup> Более подробное обсуждение этого вопроса см.: И. И. Ревзин. Модели языка (§ 52).

лингвистической традиции; д) противопоставление по методу исследования.

Первое противопоставление в его наиболее резкой форме сводится к тому, что языкознание ранее изучало язык как общественное явление, в то время как структурная лингвистика пытается изучать язык как стихийное явление, независимое от воли и действия говорящих на данном языке. Интересно, что такое противопоставление выдвигается не только идейными противниками применения математических методов в языкознании (с явной целью их компрометации<sup>8</sup>), но и некоторыми учеными, работающими в этой области<sup>9</sup>.

На первый взгляд противопоставление традиционного языкознания — науки о языке как общественном явлении — и структурной лингвистики — науки о языке как естественном явлении — представляется вполне оправданным. В то же время нельзя забывать, что уже в античных грамматиках появилась четкая система парадигм склонения и спряжения, составляющая основную часть всякого грамматического описания. Между тем только привычность этих парадигм заслоняет тот факт, что здесь мы имеем дело с глубокими структурными моделями, степень абстрактности которых весьма высока и которые описывают систему языка вне зависимости от действия и воли говорящих как естественное явление. С другой стороны, как будет показано в главе III, и отношение говорящего к языку и возникающее в связи с этим понятие нормы становится все более актуальным в структурной лингвистике.

Противопоставление по цели исследования наиболее резко сформулировано, по-видимому, С. К. Шаумяном, который противопоставил структурную лингвистику как теорию абстрактных моделей языка лингвистике традиционной как лингвистике классификационной<sup>10</sup>, или «таксономической». Такое противопоставление также имеет относительный характер. В самом деле. Именно структурная лингвистика (в особенности та ее часть, которая

<sup>8</sup> См. хотя бы: В. И. Абаев. Лингвистический модернизм как дегуманизация науки о языке. — ВЯ, 1965, № 3.

<sup>9</sup> См., например: J. P. Venzesri. *Physique et langue*. «La traduction automatique», IV, 2, 1963, стр. 32—33.

<sup>10</sup> С. К. Шаумян. Язык как семиотическая система. «Теоретические проблемы современного советского языкознания». М., 1964, стр. 48. — О том, насколько велик терминологический разнобой в этой области, свидетельствует тот факт, что Хомский почти противоположным образом противопоставлял структурную и традиционную лингвистику, ср. например: «В последующем изложении мы будем исходить из двух различных моделей порождающей грамматики. Первая, которую я называю *таксономической моделью*, является прямым развитием идей современной структурной лингвистики. Вторая, которую я называю *трансформационной моделью*, гораздо ближе к традиционной грамматике» (Н. Хомский. Логические основы лингвистической теории. Перев. с англ. «Новое в лингвистике», вып. IV. М., 1965, стр. 468).

носит название дескриптивной лингвистики) способствовала проникновению четких классификационных схем в такие области, как диалектология (ср. частые упоминания о недостаточности используемых в настоящее время в диалектологии классификационных критериев и о возможностях, открывающихся в этой области в связи с использованием структурных идей и методов). Высказывания С. К. Шаумяна можно, однако, интерпретировать, как имеющие в виду не столько противопоставление традиционной лингвистики в широком смысле и структурной лингвистики в широком смысле, сколько противопоставление теории порождающих грамматик, которую он и называет структурной лингвистикой, и описательной лингвистики, которая исходит из некоторого заданного речевого материала<sup>11</sup>. Такое разграничение, разумеется, теоретически возможно и даже подкрепляется сосюрровским противопоставлением языка (интерпретируемого как порождающий механизм) и речи (интерпретируемой как некоторое множество фраз). Однако оно резко расходится с принятой в настоящее время терминологией и ведет ко многим трудностям: например, вся теория дескриптивных моделей (как она развивается, в частности, в главах IV—VI) относится к «традиционной» проблематике, в то время как Гумбольдтово понимание языка как энергии входит в качестве составной части в структурную лингвистику.

Представляется более соответствующим сложившемуся в настоящее время словоупотреблению считать структурной лингвистикой *направление, сознательно стремящееся к четкому выявлению семиотических аспектов знака* (в абстракции от его материального воплощения), *его отношения к другим знакам в системе и в речи, т. е. структуры* (см. § 9) языков, а в связи с этим *к применению идей и методов лингвистического моделирования* (вне зависимости от того, используется ли в соответствующих моделях логико-математический аппарат и, если да, то независимо от того идет ли речь о моделях аналитических или порождающих, о синтагматических или парадигматических, о логических или статистико-вероятностных и т. п.). Что же касается традиционной лингвистики, то под ней лучше всего понимать всю совокупность идей, методов и фактических исследований, возникающих до оформления структурной лингвистики или же не использующих понятия и методы последней.

Понимаемое таким образом противопоставление удобно и для постановки вопроса об отношении структурной лингвистики к научной традиции. Как некоторые сторонники структурной лингвистики, так и в особенности многие ее противники подчерки-

---

<sup>11</sup> Именно так разъяснил С. К. Шаумян свою позицию в недавно вышедшей книге «Структурная лингвистика» (М., 1965, стр. 14—26).

вают то обстоятельство, что структурная лингвистика целиком порывает с предшествующей лингвистической традицией<sup>12</sup>.

Следует однако отметить, что вне зависимости от того, осознают ли этот факт отдельные представители структурной лингвистики (в том ее понимании, которое предложено выше) или нет, имеется непосредственная преемственность между наиболее абстрактными идеями и представлениями теории лингвистических моделей и традицией, идущей, например, в России от Фортунатова и Бодуэна через Московский лингвистический кружок и формальную грамматику 20-х годов, через поиски таких замечательных ученых, как Дурново, Петерсон, Пешковский, Щерба, Якобсон, Карцевский и многие другие.

Фактически некоторых представителей теории лингвистических моделей можно обвинить лишь в недостаточном знании существующей лингвистической традиции. Это объясняется тем, что важнейший вклад в теорию лингвистических моделей связан с деятельностью не лингвистов, а математиков. Но то, что извинительно для математиков (мы можем быть только благодарны им за то, что они сделали и делают в этой области), никак не может распространяться на лингвистов. Кроме того, в настоящее время в теории лингвистических моделей выкристаллизовались две разные науки, по разному относящиеся к лингвистической традиции:

а) математическая теория лингвистических моделей, которая строится в основном математиками и решает чисто математические проблемы (например, такие, как алгоритмическая разрешимость основных свойств языка)<sup>13</sup>;

б) лингвистическая теория лингвистических моделей, которая исследуется лингвистами и которая связана с решением таких вопросов, как соотношение модели и ее интерпретации, выявление основных лингвистических понятий, подлежащих моделированию, и основных лингвистических абстракций, использование моделей для решения проблем, возникающих в теоретической лингвистике и в ее приложениях.

Если математикам, исследующим внутреннюю структуру моделей, вообще говоря, не требуется знания лингвистических фактов (кроме некоторых уже выявленных и четко охарактеризованных исходных понятий), то, пожалуй, нет отрасли языкознания, для которой знание как можно большей части накопленного лингвистикой материала и опыта было бы столь же необходимым, как для лингвистической теории моделей, ибо последняя может развиваться в первую очередь за счет привлечения того

---

<sup>12</sup> См., например: В. Г. Адмони. Основы теории грамматики. М.—Л., 1964, стр. 6.

<sup>13</sup> В этом направлении особенно важные результаты достигнуты у нас А. В. Гладким.



материала, который лежит за пределами сегодняшнего применения структурных методов.

Существенно и другое соображение. Методы моделирования предполагают определенное упрощение языковой действительности (ср. § 4); на этом пути получены достаточно важные результаты. Так, например, в дескриптивной лингвистике достигнуты большие успехи в разработке методов дистрибутивного анализа как в фонологии, так и в морфологии и синтаксисе. Однако уже неоднократно отмечалось, что дескриптивизм есть лишь последовательное осуществление тех же позитивистских тенденций, которыми воодушевлялись младограмматики. И так же, как младограмматизм должен был быть преодолен для дальнейшего развития науки о языке, приемы и методы дескриптивной лингвистики (представляющие, разумеется, гораздо более высокую степень формализации) оказались недостаточными. Как мы говорили в § 1, на смену дескриптивизму в настоящее время приходит теория порождающих грамматик Хомского и его последователей. Общеизвестны успехи, достигнутые на этом пути, тесная связь математической теории порождающих грамматик с такими отраслями математики, как математическая логика и теория алгоритмов, которые появились именно тогда, когда возникла потребность исследовать язык математики. До последнего времени<sup>14</sup> было менее заметным то обстоятельство, с которым не может не считаться лингвистическая теория моделей, а именно — что теория порождающих грамматик возвращает нас к старой, идущей еще от Гумбольдта традиции противопоставления языка как речевого материала и языка как процесса. Больше того, вся совокупность понятий, используемая в теории порождающих грамматик, очень близка к традиционной (особенно это ясно при анализе идеи трансформационной грамматики, которая пользуется теми же смысловыми критериями, что и традиционная грамматика). Это отнюдь не означает, что теория порождающих грамматик не дает ничего нового, скорее наоборот, ее успех обеспечивается гораздо большей близостью к существующей грамматической традиции, чем это имело место в дескриптивной лингвистике.

Остается, пожалуй, лишь одно противопоставление, а именно, противопоставление по методу исследования, вернее по частным методам и средствам, по используемому подсобному аппарату. Оно действительно резко ограничивает теорию лингвистических моделей от традиционной лингвистики, ибо в последней математические методы не применялись.

---

<sup>14</sup> Теперь связь теории порождающих грамматик с учением Гумбольдта очень глубоко вскрыта самим Хомским (см.: N. Chomsky. Current issues in linguistic theory. The Hague, 1964, стр. 16—27. См. также: Н. Хомский. Логические основы лингвистической теории. Перев. с англ. «Новое в лингвистике», вып. IV, М., 1965, стр. 465—480).

Вопрос о том, почему в таких гуманитарных науках, как логика<sup>15</sup>, лингвистика и даже теория литературы<sup>16</sup>, начинают применяться математические методы и средства, достаточно разъяснен многими авторами: речь идет об использовании однозначного, с одной стороны, и чрезвычайно общего, с другой стороны, аппарата, об общей тенденции современной науки изучать не индивидуальное и неповторимое в своих объектах, а их общую структуру.

Мы бы хотели, однако, обратить внимание на один аспект использования математики, важный в первую очередь именно для лингвистики.

В той мере, в какой лингвистика имеет дело с конкретными примерами из конкретных языков, взаимопонимание между отдельными лингвистами и отдельными лингвистическими школами достигалось в первую очередь на основе приводимых примеров, понятных всем лингвистам. Этот период развития лингвистики можно было бы (разумеется, условно) уподобить тому периоду развития математики (например, индийской), когда доказательство заменялось чертежом и словами «смотри!». Мы надеемся, что не будет обидной и аналогия с методом общения лапутян у Свифта (о которых нам еще придется с уважением отзываться) — методом, состоящим в том, что произнесение слов заменялось указанием на конкретные предметы. Вне зависимости от принятия или непринятия этих аналогий следует констатировать, что там, где исследование выходило за рамки отдельных конкретных примеров, где по необходимости приходилось говорить о вещах все более абстрактных (а внутренняя структура языка по своей сути требует весьма абстрактного представления), лингвисты все меньше понимали друг друга, ибо каждый из них говорил на собственном языке. В сущности каждый лингвист по-своему понимает высказывания Гумбольдта, Гримма, Потенби и даже Соссюра, и это объясняется тем, что у лингвистов нет общего языка. Парадоксально то, что подобные ситуации много раз изучались языковедами применительно к естественным диалектам, причем выяснена и роль общего диалекта — койне — в междиалектном общении. Теория лингвистических моделей предлагает в качестве такого койне язык математики, причем именно тот диалект этого языка, который создан для обсуждения неколичественных явлений и в первую очередь для обсуждения языка самой математики и ко-

---

<sup>15</sup> Ср. например: А. Черч. Введение в математическую логику, стр. 15—16.

<sup>16</sup> См. например, содержательную статью М. Р. Майеновой «Możliwości i niebezpieczeństwa metod matematycznych w poetyce» в сборнике под характерным названием «Poetyka i matematyka» (Warszawa, 1965), ср. также статьи в сборнике под аналогичным названием «Mathematik und Dichtung» (zusammen mit Rul Gunzenhäuser herausgegeben von Helmut Kreuzer. München, 1965).

торый обладает свойствами предельной простоты, однозначности и общепонятности (для ознакомления с этим языком требуется не больше усилий, чем, например, для ознакомления с эсперанто).

Как мы увидим дальше, лингвисты, пытаясь создать свой язык для описания лингвистических оппозиций, заново строили фрагменты именно этого языка (ср. § 6).

Разумеется, для успешного использования языка математики как средства общения между лингвистами (а также и представителями смежных наук, которые гораздо лучше овладели этим языком) нужно знать не только и не столько его «словарь», т. е. совокупность наиболее элементарных математических терминов. Необходимо в первую очередь овладеть его «грамматикой», т. е. методом использования терминов (употребление словаря одного языка с грамматикой другого создает пародийный, «макаронический» стиль, весьма опасный именно для математической лингвистики).

Под «грамматикой» математического языка понимается совокупность правил строгого мышления, характерных в первую очередь для математики, необходимых, однако, в любой науке. Лингвистам это особенно необходимо. Правда, у ряда выдающихся представителей традиционной лингвистики способ мышления не очень отличался от математического (в качестве одного из последних представителей этой традиции следует назвать А. И. Смирницкого<sup>17</sup>, а еще раньше — Ф. Ф. Фортунатова, Л. В. Щербу и А. М. Пенковского, у которых мы находим логическую стройность изложения, еще не достигнутую лингвистами, применяющими логико-математические методы, — впрочем, показательно, что большинство названных ученых имели более или менее полную естественно-научную подготовку. Характерно, однако, что все эти лингвисты тяготели к конкретным рассуждениям, пользуясь в первую очередь «языком лингвистических фактов». Ясно, что они могли бы достигнуть гораздо большего, если бы у них был соответствующий язык. Пожалуй, преимущество логико-математического аппарата в том, что он позволяет каждому лингвисту, во-первых, достигать ясности и четкости, какая встречалась лишь у наиболее выдающихся мыслителей-лингвистов, и, во-вторых, говорить об абстрактных вещах, оставаясь общепонятным. При этом само собой разумеется, что математический аппарат помогает более точно формулировать мысли, но отнюдь не обеспечивает сам по себе наличие каких-нибудь мыслей.

---

<sup>17</sup> Имеются в виду работы, подготовленные к печати самим А. И. Смирницким, например: А. И. С м и р н и ц к и й. К вопросу о слове (проблема тождества слова). «Труды Ин-та языкознания», т. IV, 1954; О н ж е. Лексическое и грамматическое в слове. «Вопросы грамматического строя». М. 1955.

Тем самым еще раз подчеркивается необходимость как связи с традицией, так и продолжения конкретных исследований, ибо построение и исследование языка лингвистики никак не исчерпывает этой науки.

#### § 4. ПОНЯТИЕ МОДЕЛИ. ТРИ ТИПА ЛИНГВИСТИЧЕСКИХ МОДЕЛЕЙ

Всякое научное описание состоит в упрощении соответствующего объекта действительности, отвлечении от некоторых его сторон, изолировании его от других явлений, с которыми он может быть связан в действительности. При этом всякое движение в научном описании идет «от простого к сложному», т. е. вначале избирается максимальная степень упрощения, а затем постепенно изучаемый объект усложняется, а степень его соответствия действительности повышается. Например, физик начинает с исследования скоростей и ускорений (кинематика), причем от причин, вызывающих соответствующие ускорения, он отвлекается. Лишь после того, как изучены соответствующие законы, возникает новый аспект исследования (динамика), изучающий силы, как причины, вызывающие ускорения.

Необходимо отметить, что и теоретическая лингвистика как правило строилась по принципу упрощения, неявного моделирования, ибо каждое направление сосредоточивало свое внимание на одной какой-нибудь стороне языка, пренебрегая другими<sup>18</sup>.

Это относится тем более к конкретным исследованиям отдельных грамматических проблем. Так, анализируя три разных подхода к изучению падежа: а) чисто морфологический, б) дистрибуционалистский и в) семантический, де Гроот<sup>19</sup> показывает, что каждый из этих подходов является «симплифицирующим».

Характерны при этом следующие его высказывания: «Симплифицирующие описания неполны в том смысле, что каждое из них игнорирует, исходя из поставленной перед собой цели, факты языка, не соответствующие этой цели. Определение или описание падежей в терминах фонологических признаков игнорирует черты значения и дистрибуции падежей, хотя неявно они могут быть отражены».

Определение и описание типов словесных групп в терминах дистрибуции не включает не менее релевантные черты формы, структуры, значения и т. д. . . .

Недостатки симплифицирующих определений и описаний иллюстрируются также тем положением, что они могут быть полезны

<sup>18</sup> Это хорошо показано в работе: P. Hartmann. Modellbildungen in der Sprachwissenschaft. «Studium Generale», Jahrg. 18, Heft. 6, 1965, стр. 364—379.

<sup>19</sup> A. W. de Groot. Classification of Word Groups. «Lingua», vol. VI, 2, 1957, стр. 113—157.

как первое приближение к языкам, которыми исследователь по настоящему не владеет, поскольку он не знает точного значения слов, словесных групп и т. д. Они не могут считаться достаточными как с точки зрения синхронической, так и диахронической лингвистики — для объяснения фактов языка и языковых изменений, при обучении языку, переводе с одного языка на другой, и особенно когда имеешь дело с художественной литературой и ее интерпретацией; такие описания — это только этап на пути к структурной классификации языка в полном смысле этого слова»<sup>20</sup>.

Все это верно, если нужно подчеркнуть, что на упрощенных описаниях нельзя останавливаться, но абсолютно неверно, если имеется в виду «неполноценность» или «неуважность» упрощенных описаний, ибо без них невозможно движение науки вперед, а попытки просто перескочить через этот «скучный» этап всегда приводили к провалу.

Приходится, однако, констатировать, что тенденция к упрощению как правило вызывает раздражение у лингвистов-теоретиков. Иногда это раздражение принимало и идеологическую форму, и тогда попытки упрощения квалифицировались, например, как идеалистические. Если применительно к физике рассуждения о том, что изучение ускорений без причин, вызывающих эти ускорения, есть «бумажная физика», «никому не нужные формалистические выверты» и т. п. в лучшем случае вызвали бы смех, то в языкознании обвинения в изучении формы без содержания, звуков без значений и т. п. произносятся вполне серьезно.

Между тем языкознание столь же нуждается в упрощении, в пути от простого к сложному, как и всякая другая наука.

Метод моделей и состоит в явном введении принципа упрощения в языкознание, в четкой фиксации исходного материала и совершаемых с ним операций. Пользуясь, как было указано в § 2, языком математики (вернее очень элементарным ее подязыком), теория моделей четко фиксирует понятия и дает однозначные определения основных лингвистических объектов и операций.

Итак, моделирование — это в первую очередь упрощение. Это отнюдь не означает, что языкознание может остановиться на простых моделях, скорее наоборот, уже сейчас очевидно, что для более адекватного моделирования языка понадобятся модели принципиально гораздо более сложные, чем те, которые мы умеем строить сейчас. Вполне возможно, что эти модели воплотят в себе многие представления традиционной лингвистики. Имеется, однако, только один путь к построению этих сложных моделей будущего, а именно путь, ведущий через наиболее простые и отвлеченные модели.

---

<sup>20</sup> A. W. de Groot. Указ. соч., стр. 143.

После этих разъяснений мы попытаемся более или менее исчерпывающе описать наше понимание термина «модель» (определение термина «модель» невозможно, потому, что это понятие относится не к лингвистической теории моделей, а к мета-теории, причем к такой, которая занимается не конкретными моделями, а общими принципами применения моделей во всех науках).

Термин «модель» употребляется в лингвистике по-разному. Мы будем употреблять термин «модель» в смысле, принятом в кибернетической литературе.

Как известно, кибернетика изучает процессы, связанные с передачей и переработкой информации в механизмах и явлениях живой природы, в том числе, разумеется, и информации, передаваемой естественным языком, являющимся важнейшим средством общения людей.

В каждом процессе переработки информации можно выделить некоторую совокупность  $A$  исходных сигналов и некоторую совокупность  $B$  наблюдаемых заключительных сигналов. Задача научного описания состоит в том, чтобы объяснить, как происходит переход от  $A$  к  $B$ , каковы связи между ними. Если объекты достаточно сложны, то непосредственная регистрация зависимости между  $A$  и  $B$  часто не приводит к вскрытию каких-либо более глубоких закономерностей. Промежуточные звенья между  $A$  и  $B$  могут быть или слишком многообразными (и поэтому ускользающими от нашего наблюдения) или же просто недоступными никакому наблюдению<sup>21</sup>. Здесь особенно хорошо проявляются преимущества метода моделирования, который состоит в построении модели, т. е. некоторой системы знаков (логическое моделирование) или же некоторой системы физических объектов (физическое моделирование), обладающих следующими свойствами:

1) исходные данные модели (в кибернетике их называют «входом») соответствуют некоторой существенной части совокупности исходных объектов;

2) модель действует таким образом, что результат ее действия (в кибернетике его называют «выходом») соответствует некоторой существенной части совокупности заключительных объектов.

Описанное здесь понимание моделирования принято, например, в работах по машинному переводу и вполне согласуется с тем, как предлагает применять идеи моделирования к языку И. А. Мельчук<sup>22</sup>.

В лингвистике мы как раз имеем дело с таким объектом, в котором все промежуточные операции производятся в мозгу чело-

<sup>21</sup> В этом случае в кибернетике говорят о «черном ящике», см.: У. Росс Эшби. Введение в кибернетику. М., 1959, гл. VI.

<sup>22</sup> В кн.: О. С. Ахманова, И. А. Мельчук, Е. В. Падучева, Р. М. Фрумкина. О точных методах исследования языка. М., 1961, стр. 40—41; см. также: A. G. Oettinger. Linguistics and mathematics. «Studies presented to Joshua Whatmough on his 60 birthday», 'sGravenhage, 1957, стр. 179—186.



века и недоступны прямому наблюдению. С другой стороны, изучая язык, мы не можем отвлечься от его носителя, не можем не учитывать тех операций, которые производит с языком человек. Характерно, что в последнее время лингвисты все более остро ощущают необходимость именно такого изучения языка<sup>23</sup>.

**Примечание.** Приводимое описание модели языка сильно отличается от того, которое было дано автором ранее в книге «Модели языка».

Выбор такого описания диктуется следующими соображениями.

Описание модели, данное в книге «Модели языка» на стр. 9, может создать впечатление, что лингвистическая теория моделей мыслилась автором как аксиоматическая теория в том смысле, как этот термин понимается в математике. Между тем все дальнейшее изложение показывает, что лингвистическая теория моделей излагалась в книге не на аксиоматическом, а на содержательном уровне<sup>24</sup>. Для увязки теории лингвистических моделей с языковедческой традицией, о необходимости которой говорилось выше, именно содержательный подход к моделям и представляется наиболее плодотворным, не говоря уже о том, что на современном уровне развития теории моделей чисто аксиоматическое изложение еще просто невозможно.

Переходим теперь к понятию лингвистической модели. Речевую деятельность можно описывать с двух точек зрения: с точки зрения говорящего («синтез») и с точки зрения слушающего («анализ»). Чисто «аналитическое» или чисто «синтетическое» описание не может быть исчерпывающим, оба вида «описания одинаково существенны и должны рассматриваться как дополнительные, в смысле Н. Бора»<sup>25</sup>. Эта взаимосвязь обоих аспектов должна учитываться при описании каждого отдельного акта общения, в котором, с одной стороны, синтезируется, а с другой — анализируется данная отдельная фраза. Но после работ Хомского стало ясно, что предварительно весьма полезно описать способы, которыми может быть синтезировано или соответственно проанализи-

<sup>23</sup> Как подчеркнул недавно Р. Якобсон, «попытки построить модель языка безотносительно к говорящему или слушающему угрожают превратить язык в «схоластическую фикцию» (см.: R. Jakobson. Linguistics and communication theory. — PSAM, vol. 12, Providence, 1961, стр. 250).

<sup>24</sup> Этими соображениями автор в основном обязан доктору Карелу Берке, с которым он имел возможность во время своего пребывания в Праге обсудить ряд логических аспектов моделирования. См. также: K. Berka, P. Novák. Výklad fonologických a gramatických pojmů pomocí pojmů teorie množin. «Slovo a slovesnost», 1963, № 2, стр. 133—139.

В цитированной статье К. Берки и П. Новака предлагается следующее понимание термина «модель». Некоторую систему  $S_1$  считают моделью объекта  $S_2$ , если  $S_2$  можно гомоморфно отобразить на  $S_1$ , причем модель считается тем более адекватной, чем ближе гомоморфное соответствие к изоморфному (о понятиях гомоморфного и изоморфного отображения см. § 10). Итак, считается, что все, имеющее место в модели, имеет место и в оригинале, но обратное не обязательно. Между тем в теории лингвистических моделей мы часто должны иметь дело с идеализированными ситуациями, со свойствами, которых в действительности нет (вернее, которые намечены лишь как некоторая тенденция). Поэтому данное определение вряд ли помогает вскрыть сущность явления.

<sup>25</sup> R. Jakobson. Linguistics and communication theory, стр. 249.

зировано все множество фраз некоторого языка. Хомский подчеркнул, что построенные таким образом порождающие модели не совпадают с моделью говорящего<sup>26</sup> и соответственно обратные им модели не совпадают с моделью слушающего и это верно в том смысле, что логически возможна ситуация, когда перечислить все фразы языка удобнее одним способом, в то время как каждая фраза в действительном акте общения будет производиться иначе. Однако такая ситуация кажется маловероятной: скорее всего в речевой деятельности человека одни и те же механизмы ответственны и за перечисление всех фраз и за производство каждой конкретной фразы. Поэтому в дальнейшем (глава VII) мы будем искать модели, объединяющие черты порождения и синтеза. С другой стороны, не следует упускать из виду и серьезных логических различий описанных типов моделей.

Модель первого типа (например, синтез) имеет на входе описание определенного факта внешнего мира, т. е. совокупность десигнатов и общий смысл высказывания, а на выходе определенную фразу, т. е. последовательность слов. Если взять  $n$  таких моделей, то мы получим на выходе конечное число фраз, но мы не можем, по-видимому, получить бесконечного множества фраз, ибо для этого надо было бы уметь зафиксировать бесконечное число фактов внешнего мира.

Модель второго типа (например, порождающая грамматика) и отличается тем, что она при конечном входе порождает бесконечное множество фраз. Это достигается тем, что эта модель абстрагирована от реального содержания высказываний и учитывает лишь некоторые общие грамматические и смысловые категории.

Эта абстракция необходима для построения удобной модели, но так же необходимо помнить, что это делается лишь для упрощения исследования. Реальный механизм производства речи, описать который мы стремимся, существует, по-видимому, не столько для того, чтобы определить все множество правильно построенных фраз, сколько для того, чтобы каждый раз построить (соответственно проанализировать) языковое выражение для заданного факта внешнего мира. От того, будем ли мы учитывать в теории моделей последнее обстоятельство, существенно меняется и лингвистическая структура теории. Если ограничиться только описанием порождения множества фраз, то для каждого слова достаточно ограничиться его «формальным» значением (например, его классом эквивалентности, ср. рассуждения в § 22) — между прочим такое ограничение поля исследования было характерно для американской дескриптивной лингвистики, из которой исходил Хомский. Если же нас интересует к тому же и синтез каждого

<sup>26</sup> Н. Хомский. О понятии «правило грамматики». Перев. с англ. «Новое в лингвистике», вып. IV. М., 1965, стр. 36.

конкретного высказывания с его конкретным содержанием, то нас начинает интересовать и проблема выбора слова с данным «формальным» значением среди совокупности слов с данным «вещественным» значением (слов данной фиксированной парадигмы). Возникает проблематика сложных и многообразных отношений между вещественными и формальными значениями — давно уже занимавшая европейских лингвистов<sup>27</sup>.

Для описания этих соотношений необходимо построение специальных лингвистических моделей, существенно отличающихся от моделей первого и второго типов (порождающая грамматика и синтез).

Этот третий тип лингвистических моделей отличается от первых двух тем, что здесь на выходе появляются не конкретные фразы, а описания лингвистических фактов. Эти модели третьего типа устроены таким образом, что на входе имеется некоторая совокупность фактов, характеризующих одновременно вход и выход модели первого типа, т. е. синтеза, а именно заданы слова и их значения, фразы и их смысл. Выходом такой модели должны быть определенные категории морфологии и синтаксиса. В идеале можно было бы считать, что модели этого типа должны дать на выходе основные характеристики устройства синтеза, тогда можно было бы вслед за Хомским, который, правда, рассматривал эту задачу лишь применительно к порождающим грамматикам<sup>28</sup>, считать такую модель моделью усвоения языка. Однако пока что мы еще очень далеки от такого идеала, и модели данного типа скорее можно рассматривать как модели метаязыковой деятельности лингвиста, т. е. модели его деятельности, приводящей к установлению тех или иных категорий, связей и т. п. Моделями этого третьего типа мы и будем прежде всего заниматься в данной книге, т. е. мы все время будем учитывать интересы создания связанной с синтезом порождающей грамматики, но сначала не будем предлагать более или менее полных порождающих грамматик (не говоря уже об устройствах синтеза), а будем, пользуясь наиболее простыми порождающими грамматиками в качестве отправного пункта исследования, строить модели указанного третьего типа, т. е. модели, формализующие исследовательскую деятельность лингвиста. Это даст нам в дальнейшем возможность более глубоко исследовать некоторые аспекты построения порождающих моделей.

Таким образом, первоначальным объектом изучения в данной работе являются модели, формализующие логическую деятельность (а также, разумеется, интуицию) лингвиста, приводящую его к созданию грамматики языка, и лишь в конце книги (глава VII)

<sup>27</sup> Ср. теорию вещественно-формальных пропорциональных групп у Пауля: Г. П а у л ь. Принципы истории языка. Перев. с нем. М., 1960, стр. 128 и след.

<sup>28</sup> Н. Х о м с к и й. Логические основы лингвистической теории. «Новое в лингвистике», вып. IV. М., 1965, стр. 480—481.

будут высказаны некоторые соображения о порождающих грамматиках и моделях синтеза.

В заключение сделаем следующее терминологическое замечание. В некоторых случаях, когда не может возникнуть криво толков, мы будем употреблять термин «модель» в смысле «выход модели», «результат действия модели», например говорить «модель аморфного языка» в смысле «модель, на выходе которой получается характеристика аморфного языка», и т. п. или «модель части речи», «модель категории рода», имея в виду модель описания языка, на выходе которой появляются части речи, категории и т. п. В этом же смысле мы будем говорить о моделировании грамматических категорий, частей речи и т. п.

## § 5. ТЕОРИЯ МОДЕЛЕЙ И СТРУКТУРНАЯ ТИПОЛОГИЯ

Выбранный в данной книге угол зрения на лингвистическое моделирование связан с тем, что одной из основных ее задач является логическое обоснование предмета структурной типологии языков, в первую очередь славянских.

Дело в том, что в настоящее время сложилась достаточно парадоксальная ситуация, когда, с одной стороны, не только настойчиво подчеркивается, но и подтверждается многочисленными практическими работами<sup>29</sup> важность структурно-типологических исследований, а, с другой стороны, место структурной типологии в системе лингвистических дисциплин остается достаточно шатким и неопределенным, а вместе с тем — и это, пожалуй, более существенно — остаются невыявленными место и специфика структурно-типологического метода.

Наиболее удачное определение структурной типологии, принадлежащее Б. А. Успенскому<sup>30</sup>, все же содержит неясное указание на задачу «систематизации, инвентаризации», которая в общем-то достаточно чужда современному языкознанию. Внимательное рассмотрение анализа типологической проблематики, проведенного тем же Б. А. Успенским<sup>31</sup>, а также практической деятельности других лингвистов в этой области приводит к мысли определить предмет структурной типологии иначе, а именно:

0\*. 1. 1. Структурной типологией называется отрасль языкознания, занимающаяся исследованием моделей мета-

<sup>29</sup> См.: сб. «Структурно-типологические исследования». М., 1962 и «Исследования по структурной типологии». М., 1963.

<sup>30</sup> Ср.: «Структурную типологию языков можно определить, как систематизацию, инвентаризацию явлений разных языков по структурным признакам (т. е. признакам, существенным с точки зрения структуры данного языка)». — Б. А. Успенский. Структурная типология языков. М., 1965, стр. 10.

<sup>31</sup> Б. А. Успенский. Типологическая классификация языков как основа языковых соответствий. — ВЯ, № 6, 1961; Он же. Принципы структурной типологии. М., 1962.

языковой деятельности (моделей третьего типа, как они описаны в § 4).

Хотя такое определение и не содержит никаких указаний на сравнение языков и их систематизацию, эта задача сама собой возникает после построения некоторой модели деятельности лингвиста: поскольку конкретный языковой материал на входе этой модели является чем-то переменным, то вполне естественно систематизировать языки с точки зрения того, получается ли на выходе один и тот же результат или же разные результаты.

Тем самым основным методом структурно-типологического исследования становится метод построения моделей.

Метод моделей может оказаться, по-видимому, особенно полезным для изучения близкородственных языков, например славянских, поскольку здесь многие критерии, до сих пор рассматривавшиеся в типологии и связанные, например, со структурой слова, приводят к весьма ограниченному числу показателей, не всегда учитывающих важные расхождения в структуре языков.

Правда, именно в типологии славянских языков, как она разрабатывалась Пражским лингвистическим кружком (особенно в работах А. В. Исаченко<sup>32</sup> и В. Скалички<sup>33</sup>), выявилось важное требование изучать не целые языки, а отдельные их подсистемы, ибо, например, в славянских языках элементы флективности и агглютинативности, синтетичности и аналитичности довольно причудливо перемешаны в пределах одного и того же языка и в лучшем случае можно типологически сравнить лишь отдельные подсистемы (ср. аналитические черты в системе глагола и синтетические в системе имени большинства славянских языков). Этот круг идей будет использован нами в главе V.

Здесь необходимо, однако, сделать некоторые уточнения. Говоря о типологии в традиционном ее понимании, необходимо различать общую задачу традиционной типологии, а именно классификацию всех языков мира на основе формальных критериев, с одной стороны, и ее конкретный объект, ограниченный изучением способов выражения формальных, грамматических значений в связи с организацией слова и отчасти его отношения к предложению — с другой.

Если общая задача типологии сохраняется и в структурной типологии наших дней, то ограничение объекта исследования только явлениями, относящимися к структуре слова, в структурной типологии снимается (тенденция, проявившаяся уже в ряде работ по традиционной типологии).

<sup>32</sup> См.: А. Исаченко. Опыт типологического анализа славянских языков. «Новое в лингвистике», вып. III. М., 1963, стр. 106—120; A. I s a c h e n k o. Tense and auxiliary verbs with special reference to slavic languages. «Language», 16, 1940, стр. 189.

<sup>33</sup> В. Скаличка. О современном состоянии типологии. «Новое в лингвистике», вып. III. М., 1963, стр. 19—35.

Дело в том, что описанная выше основная задача допускает две разные интерпретации, действительно проявившиеся в ходе развития типологии: одна из них предполагает определение в минимальное число шагов типа языка — надо признать, что эта задача вполне удачно решается традиционной типологией, в особенности после важных уточнений, внесенных в эту схему Сепиром<sup>34</sup>. Правда, и здесь выяснилось, что понятие типа целесообразно понимать более гибко, перейдя от «классификационного» определения этого понятия к «ступенчатому», т. е. выясняя последовательность промежуточных ступеней, связывающих два типа между собой<sup>35</sup>. Однако общая задача все равно сводилась к определению так или иначе понимаемого «типа» языка и здесь ограничение объекта исследования в целом оправдывает себя.

Применительно к славянским языкам эта задача сводится к установлению типа данного славянского языка на основе критериев, относящихся к выделенному узкому объекту исследования. Так для чешского языка эта задача была выполнена Скаличкой<sup>36</sup>. Трудно ожидать, что для какого-нибудь другого славянского языка (за исключением болгарского и македонского) можно получить при такой интерпретации задачи типологии существенно новые результаты.

Другая интерпретация предполагает вскрытие одинаковых структурных черт двух языков, скрываемых различиями в плане выражения, и наоборот различия в формальном строении явлений двух языков при близком или даже тождественном выражении.

После этого разъяснения мы можем вернуться к существу определения 1.1. Покажем, что оно вполне соответствует подлинным представлениям типологии. Традиционная типология, задача которой состояла в определении типа языка на основе морфологической структуры слова, может быть истолкована как деятельность, в какой-то мере моделирующая работу грамматиста при описании конкретного языка; а именно при описании конкретного языка грамматист рассматривает конкретные слова, выделяя в них конкретные грамматические морфы и приписывая им конкретные значения. Если мы построим модель этой деятельности, то переменными на входе будут слова произвольного языка, а на выходе, в частности, структура слова и способы организации в ней средств выражения грамматических значений. Таким образом, традиционная типология может быть истолкована как неявное

---

<sup>34</sup> Э. Сепир. Язык. Введение в изучение речи (перев. с англ.). М.—Л., 1934.

<sup>35</sup> См.: М. И. Бурлакова, Т. М. Николаева, Д. М. Сегал, В. Н. Топоров. Структурная типология и славянское языкознание. — КСИС, вып. 33—34. «Актуальные проблемы славяноведения». М., 1961, стр. 221.

<sup>36</sup> V. Skalička. Typ češtiny. Praha, 1950.



моделирование деятельности грамматиста, связанной с изучением взаимодействия вещественных и грамматических значений в слове.

В главе V мы построим ряд частных моделей, являющихся результатом более явной модели деятельности лингвиста, связанной с исследованием взаимодействия вещественных и грамматических значений.

Однако данное нами определение позволяет расширить сферу структурной типологии, включив в нее и модели, связанные с исследованием других фактов парадигматики, например, с выявлением частей речи и других грамматических категорий (таких, как падеж, род и т. п.), а также и с исследованием фактов синтагматики. В сущности, более расширенное понимание типологии наблюдается уже достаточно давно, но связывается оно с указанным выше определением типологии как науки, задачей которой является систематизация и классификация разных языков. В конце концов такое употребление термина «типология» вполне приемлемо, если иметь в виду, что классификация языков является не самоцелью, а результатом моделирования процедур лингвистического описания. Более того, чтобы не порывать целиком с соответствующей традицией, мы будем говорить о «типологической ценности» той или иной модели, имея в виду, насколько хорошо ее применением классифицируются языки.

Заметим, что такое двойственное употребление термина «типологический» наблюдается уже достаточно давно. С одной стороны, мы имеем сочетания, вроде «типологическая ценность» с указанным выше значением, а, с другой стороны, говоря о «типологических методах» или «типологическом сравнении» (в отличие от ареального или генетического) имеют в виду сравнение не самих фактов речи, а описаний их структуры, причем существенно, что эти описания являются результатом метаязыковой деятельности.

Если сопоставить приведенное определение предмета структурной типологии с данным в § 3 описанием структурной лингвистики, то становится очевидным, что мы понимаем структурную типологию, как важную область современной структурной лингвистики, именно ту, в которой предметом изучения является не модель конкретного языка, а сравнение результатов моделирования деятельности лингвиста, связанной с построением языковых моделей для конкретных языков. Таким образом, мы предлагаем рассматривать отношение между структурным описанием конкретного языка (т. е. построением модели, на выходе которой появляются фонемы, морфемы, слова, предложения, тексты конкретного языка) и структурной типологией как параллельное отношению между грамматикой (или фонетикой) конкретного языка — русского, польского, чешского, индонезийского и т. п., с одной стороны, и общей лингвистикой, — с другой.

Представляется, что предпринятое нами переопределение задач типологии вполне соответствует духу ее развития и было уже заложено в задачах традиционной типологии.

## МНОЖЕСТВА И ОТНОШЕНИЯ. ИХ ИСПОЛЬЗОВАНИЕ В ЛИНГВИСТИКЕ

### § 6. ВВОДНЫЕ ЗАМЕЧАНИЯ

В последнее время ряд авторов<sup>1</sup> заметил, что уже достаточно давно (в особенности, после появления «Основ фонологии» Грубещкого) лингвисты начали самостоятельно разрабатывать теорию, основы которой довольно близки кругу первоначальных понятий теории множеств. И это не случайно, поскольку теория множеств занимается наиболее общими и абстрактными логическими понятиями и ее исходные понятия могут быть применены к самым разнообразным объектам. О роли понятия «множество» в науке вообще и математике в частности мы говорить не будем. Нам важно лишь применение этого понятия к объектам лингвистическим.

В лингвистике мы всегда встречаемся с некоторыми множествами объектов, например, звуков, морфем, суффиксов, слов, фраз и т. п. и с отношениями, определенными на соответствующем множестве объектов (таковы, например, оппозиции, изучаемые в фонологии, а в последнее время и в грамматике, таковы синтаксические отношения между словами в предложении и т. п.).

Более того. Если в период, предшествовавший появлению «Курса общей лингвистики» Соссюра, интересы лингвистов концентрировались вокруг выяснения свойств соответствующих объектов (например, их фонетических признаков), то Соссюр обратил внимание лингвистов именно на отношения («в каждом данном состоянии языка все покоится на отношениях»<sup>2</sup>) и с тех пор изучение отношений стало, пожалуй, основным признаком структурной лингвистики.

Но единственная разработанная до сих пор теория отношений целиком построена на алгебре множеств и является в известном смысле ее частью<sup>3</sup>. Уже одно это обстоятельство показывает, на-

<sup>1</sup> См.: J. Cantineau. Le classement logique des oppositions. «Word», vol. 11, 1955, № 1, стр. 1—9; S. Marcus. Lingvistică matematică. Modele matematice în lingvistică. București, 1963.

<sup>2</sup> Ф. де Соссюр. Курс общей лингвистики. Перев. с франц. М., 1933, стр. 121.

<sup>3</sup> Наиболее полно эта теория разработана в кн.: E. Schröder. Vorlesungen über die Algebra der Logik. Bd. III. Algebra und Logik der Re-

Сколько интересен для лингвистики теоретико-множественный, или «алгебраический», подход.

Заметим, однако, что та точка зрения, которая принята в теории отношений и состоит, как мы увидим, в том, что отношение определяется как некоторое множество пар, вполне может быть распространена и на изучение свойств.

Дело в том, что в лингвистике, по-видимому, можно сводить все рассуждения о тех или иных свойствах, например, «быть существительным», «склоняться», «спрягаться», «быть производным словом» и т. п. к рассуждениям об объеме соответствующих классов, т. е. идентифицировать некоторое свойство и класс объектов, обладающих этим свойством. Представляется, что для лингвиста то или иное свойство, например обладание категорией личности или категорией одушевленности, сводится к объему соответствующего класса (ср. гл. VI). Обычно, чтобы показать различие двух свойств, определяемых разными способами, приводят примеры, доказывающие, что существуют объекты, обладающие одним свойством и не обладающие другим. Так поступил, например, И. А. Мельчук<sup>4</sup>, когда он ввел формальные определения для разграничения свойств устойчивости и идиоматичности. Правда, имеются случаи, когда два разных свойства, например, «спрягаться» и «быть глаголом» определяют в славянских языках один и тот же класс объектов. Однако хотя здесь мы имеем дело с интуитивно разными понятиями, одно из которых связано с синтаксической ролью, а другое с системой окончаний, по-видимому, не может быть построена формализованная теория, в которой бы одно из этих свойств не определялось бы целиком через другое.

Конечно, вполне возможна лингвистическая теория, построенная на том принципе, что свойство нельзя отождествить с классом, а отношение нельзя отождествить с множеством пар. Но такая теория представляется мало продуктивной.

Во всяком случае в этой книге мы всегда будем считать, что любое свойство однозначно определяется соответствующим классом. Иначе говоря, если два свойства  $\pi_1$  и  $\pi_2$  определяют один и тот же класс лингвистических объектов, то мы будем считать, что эти свойства совпадают (так называемый «принцип объемности»<sup>5</sup>). Если вспомнить различие значения и смысла в § 1, то можно сказать, что класс соответствует значению, а свойство смыслу в избранном нами метаязыке. Выдвинутый здесь принцип объем-

lativa; см. также: А. Тарский. Введение в логику и методологию дедуктивных наук. Гл. V. «О теории отношений». Перев. с англ. М., 1948, стр. 129—162.

<sup>4</sup> И. А. Мельчук. О терминах «устойчивость» и «идиоматичность». — ВЯ, 1960, № 4, стр. 73—80.

<sup>5</sup> См.: С. А. Яновская я. [Предисловие к кн.:] Р. Карнап. Значение и необходимость. М., 1959, стр. 11.

ности можно поэтому изложить и так: для избранного нами метаязыка лингвистики — в отличие от описываемого языка — различение значения и смысла может быть снято.

## § 7. ПОНЯТИЕ МНОЖЕСТВА. ОПЕРАЦИИ НАД МНОЖЕСТВАМИ

Прежде всего остановимся на понятии «множество». Оно принадлежит к исходным понятиям, которым нельзя дать определения. В подобных ситуациях на ряде примеров объясняют, что понимают под данным исходным понятием.

Все звуки речи представляют собой некоторое множество объектов. Все слова данного языка тоже представляют собой некоторое множество объектов, некоторую совокупность. Числа 1, 2, 3, 4, 5 или 2, 4, 6, 8, 10, или 1, 4, 7, 10 представляют собой три разных множества. Для того чтобы обозначить тот факт, что множество  $M$  состоит из каких-то элементов  $e$ , употребляется запись:  $M = \{e\}$ . В нашем примере с числами:

$$M_I = \{1, 2, 3, 4, 5\};$$

$$M_{II} = \{2, 4, 6, 8, 10\};$$

$$M_{III} = \{1, 4, 7, 10\}.$$

Скобки как бы символизируют акт объединения группы объектов в некую совокупность.

Рассматриваются также множества, состоящие только из одного элемента, например,  $M_{IV} = \{1\}$ . Считается также, что множество может не содержать ни одного элемента. Это так называемое пустое множество. В дальнейшем пустое множество условно обозначается знаком  $\emptyset$ .

Таким же неопределяемым, исходным понятием является отношение принадлежности некоторого элемента  $e$  к множеству  $M$ . Если элемент  $e$  принадлежит к множеству  $M$ , то это записывается в виде  $e \in M$ , а если  $e$  не принадлежит  $M$ , то пишут  $e \notin M$ .

0. 2. 1. Говорят, что множество  $M_1$  является подмножеством множества  $M_2$ , или  $M_1$  находится в отношении включения к  $M_2$  (это обозначается при помощи  $M_1 \subseteq M_2$ ), если каждый элемент, входящий в множество  $M_1$ , входит и в множество  $M_2$ . Если  $M_1$  подмножество  $M_2$ , то  $M_2$  называется объемлющим множеством для  $M_1$ .

Считается, что пустое множество является подмножеством любого множества.

В наших примерах  $M_{IV} = \{1\}$  есть подмножество множеств  $M_I = \{1, 2, 3, 4, 5\}$  и  $M_{III} = \{1, 4, 7, 10\}$ .

Определим теперь понятие равенства двух множеств.

0. 2. 2. Говорим, что множества  $M_1$  и  $M_2$  равны и пишем  $M_1 = M_2$ , если  $M_1 \subseteq M_2$  и  $M_2 \subseteq M_1$ . В противном случае пишем  $M_1 \neq M_2$ .

Из определения 2. 1, в частности, следует, что каждое множество является подмножеством самого себя, т. е. для любого  $M$  верно  $M \subseteq M$ .

0. 2. 3. Говорят, что  $M_1$  является собственным подмножеством  $M_2$ , или строго включено в  $M_2$ , если  $M_1 \subseteq M_2$ ,  $M_1 \neq M_2$  и  $M_1$  непусто.

0. 2. 4. Объединением двух множеств  $M_1$  и  $M_2$  (обозначается  $M_1 \cup M_2$ ) называется множество, состоящее из всех элементов, которые принадлежат хотя бы одному из множеств  $M_1$  или  $M_2$ .

Объединением множеств  $M_I$  и  $M_{III}$  будет множество, состоящее из элементов 1, 2, 3, 4, 5, 7 и 10. Каждый из этих элементов принадлежит хотя бы одному из множеств  $M_I$  или  $M_{III}$ .

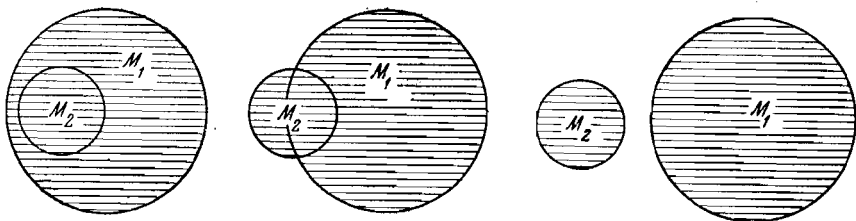
0. 2. 5. Пересечением двух множеств  $M_1$  и  $M_2$  (обозначается  $M_1 \cap M_2$ ) называется множество элементов, которые одновременно принадлежат обоим множествам  $M_1$  и  $M_2$ .

Возьмем, например, пересечение множеств  $M_{III}$  и  $M_{II}$ . По определению это должно быть множество, состоящее из всех тех элементов, которые одновременно входят в множество  $M_{II}$  и множество  $M_{III}$ . Легко видеть, что  $M_{III} \cap M_{II} = \{4, 10\}$ .

Пересечение  $M_{IV} \cap M_{II}$  не содержит ни одного элемента и поэтому является пустым множеством, т. е.  $M_{IV} \cap M_{II} = \emptyset$ .

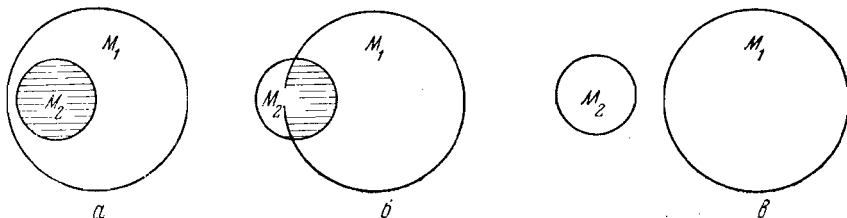
Легко видеть, что в случае, когда  $M_1$  есть подмножество  $M_2$ , имеют место следующие равенства:  $M_1 \cup M_2 = M_2$  и  $M_1 \cap M_2 = M_1$ .

Операции над множествами можно представить следующими схемами.



Здесь заштрихованная часть представляет объединение  $M_1 \cup M_2$  в трех случаях: а) когда  $M_2$  есть подмножество  $M_1$ ; б) когда  $M_1$  и  $M_2$  пересекаются, но ни одно из них не есть подмножество другого и в) когда  $M_1$  не пересекается с  $M_2$ .

В схеме 2.2. в тех же трех случаях заштрихованная часть представляет пересечение  $M_1 \cap M_2$  (в случае в  $M_1 \cap M_2 = \emptyset$ ):



В случае, когда зафиксировано некоторое исходное множество  $U$  (например, множество всех звуков русского языка), бывает полезным и следующее понятие.

**0.2.6.** Дополнением множества  $M$  до множества  $U$  (обозначается  $\bar{M}$ ) называется множество тех элементов  $U$ , которые не принадлежат  $M$ .

Например, если исходным множеством является множество звуков русского языка, причем каждый звук либо гласный, либо согласный и если  $M$  есть множество всех гласных звуков-типов — пусть это будет для простоты:  $\{i, y, o, u, e, \epsilon, a\}$ , то множество  $\bar{M}$  есть множество всех согласных.

При характеристике исходных теоретико-множественных понятий всегда различают конечные и бесконечные множества. Большинство множеств, рассматриваемых в лингвистике, например, множества звуков, фонетических признаков, слов, морфем, грамматических категорий, суть конечные множества. Однако в лингвистике мы встречаемся и с бесконечными множествами: таково, например, множество всех возможных текстов на данном языке. Говоря о всех возможных текстах, мы вовсе не имеем в виду, что нам задано сразу все это бесконечное множество и мы можем производить с ним какие-то операции — подобная точка зрения,

плодотворная в некоторых отраслях математики и связанная с так называемой абстракцией «актуальной бесконечности», в лингвистике, по-видимому, неприемлема. Мы имеем в виду, что какое бы конечное множество из  $n$  текстов мы ни взяли, всегда имеется возможность построить еще один  $(n+1)$ -ый текст, отличный от них (это так называемая абстракция «потенциальной оусуществимости»).

Мы закончили характеристику исходных теоретико-множественных понятий. Теперь мы введем одно понятие, которое неоднократно будет использовано нами в дальнейшем.

**0.2.7.** Пусть нам дано некоторое множество  $M$  и некоторая совокупность его непустых подмножеств  $M_1, M_2, M_3, \dots, M_n$ . Эта совокупность называется разбиением множества  $M$  (а соответствующие подмножества называются классами этого разбиения), если каждый элемент множества  $M$  входит в одно и только одно множество  $M_i$ .

Иначе говоря, требуется, чтобы все подмножества  $M_i$  были непустыми, чтобы никакие два разных подмножества  $M_i$  и  $M_j$  не пересекались и чтобы каждый элемент, принадлежащий  $M$ , входил хотя бы в одно из подмножеств  $M_i$ . Возьмем некоторое множество и все его разбиения  $B_1, \dots, B_m$ .

**0.2.8.** Разбиение  $B_1$  называется укрупнением разбиения  $B_2$ , если любые два элемента, входящие в один класс разбиения  $B_2$ , входят в один класс разбиения  $B_1$ . Разбиение  $B_2$  называется правильным укрупнением разбиения  $B_1$ , если оно есть укрупнение  $B_1$  и  $B_1 \neq B_2$ .

Возьмем, например, множество слов русского языка (в исходной форме, т. е. глаголы в неопределенной форме, существительные в именительном падеже и т. п.). Пусть разбиение  $B_1$  есть разбиение, в один класс которого попадают все те и только те слова, которые оканчиваются на одинаковую букву, а разбиение  $B_2$  есть разбиение, в один класс которого попадают все слова, имеющие одинаковое окончание (суффиксы или флексии). Тогда два слова, попадающие в один класс разбиения  $B_2$ , попадут и в один класс разбиения  $B_1$ , но обратное неверно (*старость* и *плакать* попадут в один класс разбиения  $B_1$ , но в разные классы разбиения  $B_2$ ). Как и следовало ожидать, разбиение по последней букве есть укрупнение разбиения по суффиксу.

## § 8. ПОНЯТИЕ БИНАРНОГО ОТНОШЕНИЯ

В разных областях жизни, в математике, в лингвистике мы имеем дело с высказываниями об отношениях, например,

- 1)  $a$  женат на  $b$
- 1а)  $a$  ist verheiratet mit  $b$
- 2)  $a$  старше  $b$
- 3)  $a$  родитель  $b$
- 4) слово  $a$  рифмуется со словом  $b$
- 5) слово  $a$  согласуется со словом  $b$
- 6) слово  $a$  имеет тот же корень, что и слово  $b$
- 7) слово  $a$  производное от  $b$
- 8) слово  $a$  синонимично слову  $b$

Во всех этих и подобных случаях мы имеем дело с одной и той же исходной ситуацией. Имеется некоторое множество объектов  $M$ , например, множество людей в случаях 1, 1а, 2, 3, или множество слов языка в случаях 4—8. Мы рассматриваем произвольные пары объектов этого исходного множества, причем для некоторых пар данное отношение имеет место, а для других не имеет. Множество пар, для которых соответствующее высказывание имеет место, однозначно определяет отношение. Иначе говоря, здесь действует тот принцип объемности, о котором мы говорили в § 6: если два каких-нибудь отношения определяют одно и то же множество пар, то мы будем считать их совпадающими. Это приводит к следующей процедуре.

Рассмотрим некоторое множество  $M$ . Пусть для простоты множество состоит всего из трех элементов  $M_1 = \{a, b, c\}$ . Мы можем интерпретировать его как совокупность трех лиц, например  $a = \text{Вадим}$ ,  $b = \text{Пульхерия}$ ,  $c = \text{Мария}$ , или же как совокупность трех слов:  $a = \text{солнце}$ ,  $b = \text{сонное}$ ,  $c = \text{светило}$ . Составим все возможные пары элементов, принадлежащих некоторому множеству, например,  $M_1$ , причем пары, различающиеся порядком элементов, мы будем считать разными, и, кроме того, в число наших пар включим и те, в которых на первом и на втором месте стоит тот же самый элемент.

В нашем случае множество всех пар состоит из 9 членов:  $\langle a, a \rangle$ ,  $\langle a, b \rangle$ ,  $\langle a, c \rangle$ ,  $\langle b, a \rangle$ ,  $\langle b, b \rangle$ ,  $\langle b, c \rangle$ ,  $\langle c, a \rangle$ ,  $\langle c, b \rangle$ ,  $\langle c, c \rangle$ .

0.2.9. Обозначим через  $M^2$  множество всех пар, составленных из  $M$ . Будем говорить, что над исходным множеством  $M$  определено б и н а р и о е о т н о ш е н и е  $R$ , если нам задана формальная процедура, выделяющая некоторое множество пар элементов  $R$  такое, что  $R \subseteq M^2$ .

Рассмотрим, например, множество из трех лиц: *Вадим*, *Пульхерия*, *Мария*. Если *Вадим женат на Пульхерии*, то это отношение можно зафиксировать как упорядоченную пару  $\langle \text{Вадим}, \text{Пульхерия} \rangle$ . Отношение *Vadim ist mit Pulcheria verheiratet* фиксируется двумя упорядоченными парами:  $\langle \text{Vadim}, \text{Pulcheria} \rangle$ ,  $\langle \text{Pulcheria}, \text{Vadim} \rangle$ . Это различие связано с тем, что отношение *Пульхерия жената на Вадиме* не имеет места, в то время как имеет место отношение *Pulcheria ist mit Vadim verheiratet*. Пусть теперь *Мария мать Вадима*, а *Вадим старше Пульхерии*. Тогда отношение « $a$  старше  $b$ » будет представлено парами:  $\langle \text{Мария}, \text{Вадим} \rangle$ ,  $\langle \text{Вадим}, \text{Пульхерия} \rangle$ ,  $\langle \text{Мария}, \text{Пульхерия} \rangle$ , а отношение « $a$  родитель  $b$ » будет представлено парами  $\langle \text{Мария}, \text{Вадим} \rangle$ . Аналогичные примеры можно привести и для отношений на множестве слов, но поскольку сам принцип, по-видимому, уже ясен, то мы в дальнейшем для краткости ограничимся абстрактными, неинтерпретированными отношениями на множестве  $M_1 = \{a, b, c\}$ .

Бинарное отношение удобно задавать в виде матрицы (таблицы), где каждая единичка обозначает наличие пары, первый элемент которой берется из соответствующей строки, а второй из соответствующего столбца. Например, бинарное отношение  $R_1$  над множеством  $M_1 = \{a, b, c\}$  и состоящее из пар  $\langle a, a \rangle$ ,  $\langle a, b \rangle$ ,  $\langle b, a \rangle$ ,  $\langle c, c \rangle$ ,  $\langle b, b \rangle$ ,  $\langle b, c \rangle$  может быть представлено следующей матрицей:

$R_1$	$a$	$b$	$c$
$a$	1	1	0
$b$	1	1	1
$c$	0	0	1

Тот факт, что пара  $\langle x, y \rangle$  входит в бинарное отношение  $R$  мы будем обозначать двумя следующими способами:  $xRy$  или  $\langle x, y \rangle \in R$ .

В дальнейшем вместо «бинарного отношения» мы будем писать просто «отношение», поскольку других отношений мы здесь касаться не будем<sup>6</sup>.

<sup>6</sup> Аналогично бинарному отношению можно было бы определить «тернарное отношение» как подмножество множества всех упорядоченных троек и т. п. Заметим, что в последнее время делаются попытки изложить некоторые лингвистические проблемы в терминах теории тернарных отношений, например: L. Nebeský, P. Sgall. *Vztah formy a funkce v jazyce*. «Slovo a slovesnost», 1962, № 3, стр. 174—189.

Изложим теперь некоторые понятия теории отношений.

**0.2.10.** Отношение называется **дополнительным** к отношению  $R$  и обозначается  $\bar{R}$ , если в него входят те и только те пары, которые не входят в  $R$ , т. е.  $x\bar{R}y$ , если неверно  $xRy$ . Так дополнительным к  $R_1$  из нашего примера будет отношение  $\bar{R}_1$ , которому соответствует матрица:

$\bar{R}_1$	$a$	$b$	$c$
$a$	0	0	1
$b$	0	0	0
$c$	1	1	0

**0.2.11.** Отношение  $R$  называется **рефлексивным**, если для любого элемента  $x$ , входящего в исходное множество  $M$  верно  $xRx$ , т. е. что пара  $\langle x, x \rangle$  входит в  $R$ . Примером рефлексивного отношения служит  $R_1$ . Отношение называется **антирефлексивным**, если дополнительное к нему отношение рефлексивно. Приведем теперь пример отношения, которое не является ни рефлексивным, ни антирефлексивным.

$R_2$	$a$	$b$	$c$
$a$	0	1	0
$b$	1	0	0
$c$	0	0	1

$\bar{R}_2$	$a$	$b$	$c$
$a$	1	0	1
$b$	0	1	1
$c$	1	1	0

**0.2.12.** Отношение  $R$  называется **симметричным**, если для любых  $x$  и  $y$ , входящих в исходное множество  $M$ , из условия  $xRy$  (пара  $\langle x, y \rangle$  входит в  $R$ ) следует  $yRx$  (пара  $\langle y, x \rangle$  входит в  $R$ ). Примерами симметричных отношений служат  $R_2$  и  $\bar{R}_2$ . Отношение называется **антисимметричным**, если для любых  $x$  и  $y$ , входящих в исходное множество  $M$ , из условия: пара  $\langle x, y \rangle$  входит в  $R$ , следует, что пара  $\langle y, x \rangle$  не входит в  $R$ . В качестве примера антисимметричного отношения можно привести  $R_3$ .

$R_3$	$a$	$b$	$c$
$a$	0	1	1
$b$	0	0	1
$c$	0	0	0



0.2.13. Отношение называется транзитивным, если для любых элементов  $x, y$  и  $z$ , входящих в исходное множество, из условия, что верно  $xRy$  и  $yRz$  следует  $xRz$ . Отношения  $R_1, \bar{R}_1, R_2$  и  $\bar{R}_2$ , как легко проверить, нетранзитивны. Отношение  $R_3$  транзитивно.

0.2.14. Если некоторое отношение  $R$  является одновременно рефлексивным, симметричным и транзитивным, то оно называется отношением типа эквивалентности (если  $R$  отношение типа эквивалентности и имеет место  $xRy$ , то этот факт мы будем обозначать так:  $x \sim y$ , а в случае, когда  $R$  фиксировано и не может возникнуть кривотолков, будем просто писать  $x \sim y$ ). Приведем пример отношения типа эквивалентности над  $M_1$ :

$R_4$	$a$	$b$	$c$
$a$	1	0	0
$b$	0	1	1
$c$	0	1	1

Имеет место следующее важное утверждение.

Т. 2.1. Каждое разбиение какого-нибудь множества  $M$  на непересекающиеся подмножества определяет между элементами множества  $M$  некоторое отношение типа эквивалентности. Обратное: каждое отношение типа эквивалентности над множеством  $M$  определяет разбиение  $M$ , причем в каждый класс входят эквивалентные между собой элементы.

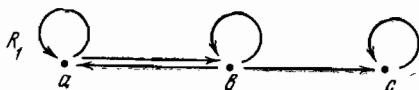
Так отношение  $R_4$  разбивает  $M$  на два подмножества: 1)  $\{a\}$  и 2)  $\{b, c\}$ .

Задача разбиения множества элементов на непересекающиеся подмножества (или классы) возникает в лингвистике довольно часто. Смысл приведенной теоремы в том, что для решения такой задачи достаточно проверить, обладает ли то отношение, на основе которого проводится классификация, свойствами эквивалентности (некоторые лингвистические примеры будут приведены ниже в § 13).

0.2.15. Всякое множество  $M$  с заданным на нем отношением  $R$  называется графом  $(M, R)$ . Любой  $x \in M$  называется в этом случае вершиной графа, а всякая пара  $\langle x, y \rangle$  такая, что имеет место  $xRy$ , называется ребром графа.

Термин «граф» имеет своим источником следующий графический метод изображения. Изобразим все вершины графа точками на плоскости, а каждое ребро  $\langle x, y \rangle$  изобразим стрелкой, соединяющей  $x$  и  $y$  и направленной от  $x$  к  $y$ .

В частности, отношение  $R_1$  из нашего примера может быть представлено следующей схемой.



Остальные отношения, рассмотренные в нашем примере, получают схемы, представленные на стр. 41.

0. 2. 16. Граф  $(M_1, R_1)$  является подграфом графа  $(M_2, R_2)$ , если  $M_1 \subseteq M_2$  и любая пара  $\langle x, y \rangle$ , входящая в  $R_1$ , одновременно входит и в  $R_2$ .

Пусть у нас имеются некоторые отношения  $R_1, R_2, \dots, R_n$ . Исходя из этих отношений, мы можем получать новые отношения. С одним таким способом мы уже познакомились: это образование дополнительного отношения. В дальнейшем нам понадобятся еще следующие операции.

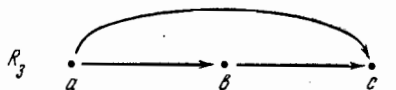
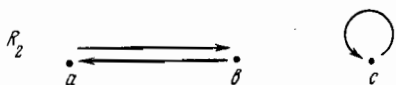
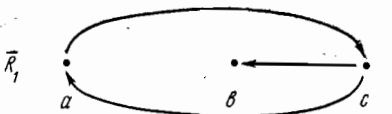
0. 2. 17. Объединением двух отношений  $R_1$  и  $R_2$  (обозначается  $R_1 \cup R_2$ ) называется отношение, состоящее из всех тех пар  $\langle x, y \rangle$ , которые входят хотя бы в одно из отношений  $R_1$  или  $R_2$ .

0. 2. 18. Транзитивным замыканием отношения  $R$  (обозначим его  $R^+$ ) называется отношение, в которое входят все такие пары  $\langle x, y \rangle$ , для которых в исходном множестве  $M$  найдутся элементы  $z_1, \dots, z_n$  такие, что

- а)  $z_1 = x$ ,
- б)  $z_n = y$ ,
- в) для любого  $i$  ( $1 \leq i < n$ ) имеет место  $z_i R z_{i+1}$ .

0. 2. 19. Симметризацией отношения  $R$  (обозначим его  $R^c$ ) называется отношение, в которое входят все пары  $\langle y, x \rangle$  и все пары  $\langle x, y \rangle$ , для которых имело место  $x R y$ .

0. 2. 20. Возьмем симметризацию некоторого отношения  $R$  и получим затем транзитивное замыкание отношения



нения  $R^c$ , т. е. новое отношение  $(R^c)^+$ . Граф  $(M, R)$  называется связным, если для любых  $x$  и  $y$  из  $M$  имеет место  $x (R^c)^+ y$ .

На схеме хорошо видно, что связными являются графы для отношений  $R_1, R_1, R_2$  и  $R_3$ . В то же время графы для отношений  $R_2$  и  $R_4$  не являются связными. При графическом представлении смысл связности весьма нагляден: граф связан, если мы можем (вне зависимости от направлений стрелки — в этом смысле симметризация) от любого элемента  $x$  дойти до любого элемента  $y$ , двигаясь по стрелкам.

0. 2. 21. Назовем циклом графа  $(M, R)$  последовательность его вершин  $x_1 \dots x_n$ , обладающих теми свойствами, что  $x_i R x_{i+1}$  и  $x_1 = x_n$ .

Так, граф отношения  $R_1$  имеет три цикла, ибо имеет место  $a R_1 a$ ,  $b R_1 b$  и  $c R_1 c$ . Из этого примера ясно, что любое отношение, не являющееся антирефлексивным, дает граф с циклами. Кроме того, здесь имеется цикл  $a R_1 b$  и  $b R_1 a$ . Ясно поэтому, что не антисимметричное отношение также содержит циклы.

0. 2. 22. Конечный граф  $(M, R)$  называется деревом с корнем  $x_1$ , если:

- а) он является связным и не содержит циклов;
- б) ни для одного  $x$  не имеет места  $x R x_1$ .
- в) для любого  $x$  такого, что  $x \neq x_1$ , существует одно и только одно  $y$  такое, что  $y R x$ .

**Примечание.** Мы употребляем термин *дерево* в более узком смысле, принятом применительно к лингвистическим задачам. В теории графов термин *дерево* употребляется во всех случаях, когда выполняется условие а), а для нашего понятия употребляется термин «прадерево с корнем  $x_1$ »<sup>7</sup>. Между прочим, как показано, например, в цитированной работе К. Бержа, из свойств б) и в) свойство связности можно вывести. Мы упомянули о нем, во-первых, ввиду его важности и, во-вторых, потому, что так легче увидеть разницу между «деревом» в обычном более широком смысле и «деревом» в нашем употреблении.

Примером дерева, хорошо знакомого каждому лингвисту, является генеологическое дерево родства индоевропейских языков (с праиндоевропейским в качестве корня), предложенное Шлейхером.

В лингвистической теории важную роль играет один частный случай дерева, а именно отношение линейного порядка<sup>8</sup>, определенное на конечном множестве.

**0. 2. 23.** Отношение  $R$  называется отношением линейного порядка, если оно

а) антирефлексивно и антисимметрично;

б) для любых  $x$  и  $y$  таких, что  $x \neq y$  верно  $xRy$  или  $yRx$ .

Отношением линейного порядка является, например, отношение: слово  $a$  находится в тексте правее слова  $b$  (корнем является самое левое слово, т. е. первое слово текста).

## § 9. ПОНЯТИЕ ОТОБРАЖЕНИЯ

Весьма важным классом отношений являются функциональные отношения, или отображения.

**0. 2. 24.** Отношение  $R$  называется отображением, или функциональным отношением, если для любых  $x, y$  из того, что имеет место  $xRy$  и  $zRy$ , следует  $x=z$ .

В наших примерах функциональным отношением является отношение  $R_2$ .

Здесь необходимо одно небольшое разъяснение. Данное определение предполагает как будто бы, что отображение обязательно определяется в пределах одного множества, в то время как обычно под функциональным отношением имеется в виду отношение между элементами двух множеств  $M_1$  и  $M_2$ .

Но все дело в том, что исходное множество  $M$ , на котором определяются все отношения, вполне может быть в частном случае и объединением двух множеств  $M = M_1 \cup M_2$ . Если все первые члены пар вида  $\langle x, y \rangle$  принадлежат к  $M_1$ , а все вторые члены к  $M_2$ , то данное нами определение требует, чтобы между  $M_1$  и  $M_2$  было установлено такое соответствие, когда какому-либо элементу из  $M_2$  не может соответствовать более одного элемента из  $M_1$ .

Если теперь на  $M_1 \cup M_2$  определено функциональное отображение  $F$ , т. е. множество пар вида  $\langle x, y \rangle$ , где  $x \in M_1, y \in M_2$ , то естественно в случае, когда имеет место  $xFy$  (т. е.  $\langle x, y \rangle \in F$ ), писать  $x = F(y)$ , причем эта запись в силу определения однозначно определяет отношение.

Приведем теперь примеры отображений.

Всякий, работающий с двуязычным словарем, знает, что для того, чтобы найти перевод какого-нибудь слова (точнее словоформы), встретившегося в тексте, например *стол* или *идешь*, нужно прежде всего знать исходную

<sup>7</sup> См.: К. Берж. Теория графов и ее применения. Перев. с франц. М., 1962, гл. 16.

<sup>8</sup> На важность отношения линейной упорядоченности («протяженности в одном измерении») указал еще Соссюр в своем «Курсе общей лингвистики».

форму этого слова, например, форму именительного падежа единственного числа существительного или инфинитива глагола. Таким образом, пользование словарем предполагает, что установлено соответствие между множеством словоформ и множеством исходных форм, причем это отношение соответствия является отображением, поскольку каждой словоформе поставлена в соответствие одна и только одна исходная форма. Это отображение можно представить как ряд пар вида:

$идти = F(иду)$	$стол = F(стол)$
$идти = F(идешь)$	$стол = F(стола)$
.....	.....
$идти = F(идут)$	$стол = F(столов)$
$идти = F(идти)$	и т. п.

Другой вид функционального отношения, очень часто используемого в лингвистике, это отношение между звуками (звуками-типами) и совокупностями признаков. Каждому звуку ставится в соответствие одна и только одна совокупность фонологических признаков.

Отношение между звуками и буквами в русском письме (или, наоборот, между буквами и звуками) не является отображением, ибо не только одной букве могут соответствовать разные звуки (например, букве *ч* соответствует *ʃ* и *ʧ*), но одному и тому же звуку, например *j*, могут соответствовать разные буквы. В сербской орфографии каждому звуку соответствует одна и только одна буква и здесь мы имеем дело с отображением в строгом смысле слова (этот пример показывает лингвистическую важность понятия отображения, а именно это тот идеал соответствия, к которому всегда стремятся в языке).

0. 2. 25. Пусть определено отображение  $F$ . Совокупность  $x$ , для которых существует  $y$  такое, что имеет место  $xFu$  (т. е. при нашем способе определения отображения — множество  $M_1$ ) называется областью значения и  $y$  отображения, а совокупность  $y$ , для которых существует  $x$ , такое, что  $xFu$  (т. е. множество  $M_2$ ) называется областью определения отображения. Элемент  $x$  такой, что  $x=F(y)$ , называется образом элемента  $y$ , а элемент  $y$  называется прообразом элемента  $x$ . Множество всех  $y$  таких, что  $x=F(y)$  называется полным прообразом элемента  $x$ .

Рассмотрим теперь некоторые специальные виды отображений.

0. 2. 26. Пусть задано множество  $M_1$  с некоторым отношением  $P$  и множество  $M_2$  с некоторым отношением  $Q$  и, кроме того, задано отображение  $F$ , ставящее в соответствие элементам из  $M_1$  элементы из  $M_2$ . Отображение  $F$  называется гомоморфным отображением (или гомоморфизмом), если для любых  $x$  и  $y$  таких, что  $xPy$  верно, что  $(F(x))Q(F(y))$ , где  $F(x)$  и  $F(y)$  элементы, соответствующие  $x$  и  $y$  при отображении  $F$ .

То, что гомоморфизм на самом деле вполне естественное для лингвистики понятие, показывает следующий пример. Вновь рассмотрим множество всех словоформ и определим на нем отношение: словоформа  $a$  входит в ту же парадигму словоизменения<sup>9</sup>, что и словоформа  $b$  (иначе говоря,  $a$  и  $b$  являются формами одного слова) — заметим, что это отношение типа эквивалентности; рассмотрим также множество исходных («словарных» форм) и определим на нем отношение  $Q$ : исходная форма  $a$  начинается с той же буквы, что и исходная форма  $b$  (это также отношение типа эквивалентности). Определим наконец отображение  $F$  так, чтобы все словоформы отображались в свою исходную форму. Легко проверить, что если две словоформы, например *сделанный* и *делаю*, входят в одну парадигму, то они соответствуют исходной форме, начинающейся с той же буквы (в данном случае той же самой).

<sup>9</sup> Это отношение в дальнейшем рассматривается нами весьма подробно как отношение принадлежности к одной парадигме.

Таким образом, это отображение есть гомоморфизм. Содержательный смысл понятия гомоморфизма состоит в укрупнении (и тем самым упрощении) объекта исследования при сохранении некоторых существенных отношений.

Очень важными являются для лингвистики и следующие понятия.

0. 2. 27. Отображение  $F$ , определенное на  $M_1$  и имеющее  $M_2$  областью значений, называется *взаимнооднозначным*, если каждому элементу из  $M_2$  соответствует ровно один элемент из  $M_1$  (иначе говоря, взаимнооднозначным отображением называется такое отношение, когда каждому элементу из  $M_1$  соответствует ровно один элемент из  $M_2$  и каждому элементу из  $M_2$  соответствует ровно один элемент из  $M_1$ ). Отображение, являющееся взаимнооднозначным и гомоморфным, называется *изоморфным* отображением, или *изоморфизмом*.

В этом случае мы будем говорить, что графы  $(M_1, P)$  и  $(M_2, Q)$  — где  $P$  и  $Q$  отношения, относительно которых установлен гомоморфизм — *изоморфны*.

Отношение изоморфизма на множестве всех графов есть отношение типа эквивалентности и оно разбивает это множество на классы изоморфных графов, которые мы будем называть *структурами*.

Таким образом, структура в нашем понимании это граф, природа элементов которого для нас несущественна, так же как и конкретное отношение, определенное на нем. Нас интересует лишь число элементов и абстрактные свойства соответствующего отношения. Теперь мы можем уточнить понимание термина «структурная лингвистика», предложенное в § 3, а именно структурная лингвистика есть отрасль языкознания, изучающая структуру лингвистических явлений, т. е. изучающая лингвистические явления с точностью до изоморфизма. Ее интересует то общее, что есть в материально разных лингвистических объектах как в парадигматическом, так и в синтагматическом плане. Собственно говоря, в какой-то мере так поступало и традиционное языкознание, отвлекаясь, например, от различий устной и письменной речи. Структурная лингвистика требует лишь более четкого выделения уровня, на котором соответствующие явления действительно относятся к одной и той же структуре. Вернемся к примеру, приведенному выше.

Соответствие между буквами и звуками в сербской орфографии может считаться взаимнооднозначным, но его нельзя называть изоморфизмом до тех пор, пока не установлено, о сохранении каких отношений может идти речь.

Изоморфизм чрезвычайно важен в качестве некоторого идеального (но не достигаемого на практике) образца например, во всех случаях, когда речь идет о переводе, ср.

<i>свеча горела на столе</i>	Die Kerze brannte auf dem Tisch
<i>свеча горела (Пастернак)</i>	Die Kerze brannte (немецкий перевод),

где не только установлено взаимно-однозначное соответствие между словами, но сохраняются отношение порядка, отношение синтаксической зависимости.

## § 10. ДЕРЕВО ЛОГИЧЕСКИХ ВОЗМОЖНОСТЕЙ И ПОНЯТИЕ АЛГОРИТМА

В целом ряде лингвистических задач мы встречаемся с ситуацией, когда одновременно задается множество лингвистических объектов (звуков, слов, фраз и т. п.) и некоторое множество свойств или признаков этих объектов, причем современная лингвистика умеет как правило выбирать такие признаки, что каждый из объектов либо обладает, либо не обладает соответствующим признаком или во всяком случае каждый признак принимает ровно два значения (принцип бинаризма).

В частности, для фонологии Р. Якобсоном и его сотрудниками предложены признаки, позволяющие описывать фонемный состав любого языка.

Это следующие признаки <sup>10</sup>:

1. гласность—негласность
2. согласность—несогласность
3. компактность—некомпактность
4. диффузность—недиффузность
5. периферийность—непериферийность
6. непрерывность—прерывность
7. дизность—недизность
8. напряженность—ненапряженность
9. назальность—неназальность
10. яркость—тусклость
11. звонкость—глухость
12. долгота—краткость
13. палатальность—непалатальность
14. бемольность—небемольность

Эта система признаков удобна в первую очередь тем, что в ней одни и те же признаки используются для описания как гласных так и согласных звуков. Нас здесь интересует не данная система как таковая, а общая ситуация, которая в случае звуков языка представляется наиболее простой.

Итак, пусть имеется множество объектов и ряд признаков  $\pi_1, \pi_2, \dots, \pi_n$ , дифференциальных в том смысле, что имеются хотя бы два объекта данного множества, которые различаются по этому признаку. Значения каждого признака (в простейшем случае их два: объект обладает или не обладает данным признаком) разбивают все множество на ряд подмножеств относительно значений этого признака (в простейшем случае на два подмножества, а именно объекты, обладающие и не обладающие данным признаком). Система соответствующих подмножеств для каждого признака однозначно определяет данный признак.

0. 2. 28. Разделим исходное множество  $M$  на два подмножества  $M_0$  и  $M_1$  в соответствии с тем, обладают или не обладают соответствующие объекты первым признаком, и назовем их подмножествами 1-го ранга. Каждое из подмножеств 1-го ранга разобьем на два по второму признаку, получим четыре подмножества второго ранга  $M_{00}, M_{01}, M_{10}$  и  $M_{11}$ . Если одно из этих подмножеств пусто, то мы будем называть его пустой клеткой. В общем случае пусть мы имеем  $2^i$  подмножеств  $i$ -го ранга и имеется еще один  $(i+1)$ -ый дифференциальный признак. Тогда мы разобьем каждое из подмножеств  $i$ -го ранга на два подмножества  $(i+1)$ -го ранга по  $(i+1)$ -му признаку и, если одно из них пусто, будем называть его пустой клеткой. Получившуюся совокупность непустых подмножеств с отношением строгого включения ( $M_1$  строго включено в  $M_2$ , если  $M_1 \subseteq M_2$  и  $M_1 \neq M_2$ ) назовем деревом логических возможностей.

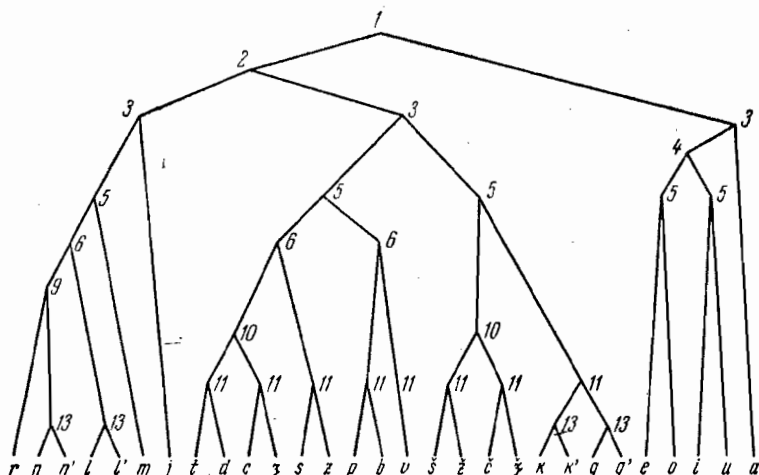
Примечания. 1) Легко проверить, что отношение строгого включения на получившейся совокупности подмножеств удовлетворяет определению 2.22 (корнем дерева является исходное множество). 2) Количество пустых клеток отражает степень несистемности данного множества объектов относительно выбранного способа упорядочения. Если число пустых клеток намного превышает число объектов, то данную со-

<sup>10</sup> Мы далее остановимся на том варианте системы дифференциальных признаков, который принят Сектором структурной типологии славянских языков Института славяноведения и отражен в работе: М. И. Лекомцева, Д. М. Сегал, Т. М. Судник, С. М. Шур. Опыт построения фонологической типологии близкородственных языков. «Славянское языкознание. Доклады советской делегации. V Международный съезд славистов». М., 1963, стр. 423—476.

вокущность лингвистических объектов нецелесообразно называть лингвистической (фонологической, лексической, грамматической и т. п.) системой<sup>11</sup>.

Заметим теперь, что схемы идентификации фонем в терминах дифференциальных признаков представляют собой ни что иное как деревья логических возможностей.

Рассмотрим, например, схему идентификации фонем македонского языка и юго-западных болгарских говоров.



Каждый узел однозначно определяет на этой схеме определенное подмножество фонем, а подчиненность узлов символизирует строгое включение. В цитированной работе достаточно хорошо показана ценность подобных схем для типологии славянских языков.

Мы бы хотели, однако, отметить, что подобная же ситуация нам встречается и на других уровнях языка, например в грамматике.

Рассмотрим следующие примеры<sup>12</sup>. Пусть нам дано множество словоформ существительных некоторого языка и нам надо определить род произвольного существительного.

Посмотрим, на какие подмножества надо делить все существительные чешского языка, с одной стороны, и словацкого языка, с другой (эти языки взяты потому, что генетически они входят в одну группу, а по структуре грамматических родов демонстрируют ряд резких типологических различий).

В чешском языке надо проверять следующие свойства:

1. Оканчивается ли И, ед. ч. на *-i* или не *-i*?
2. Оканчивается ли И, ед. ч. на *-e* или не *-e*?

<sup>11</sup> Соображения по этому поводу и соответствующие количественные показатели изложены: И. А. М е л ь ч у к. К вопросу о термине «система» в лингвистике. «Zeichen und System in der Sprache», Bd. II. Berlin, 1962; О н ж е. О стандартной форме и количественных характеристиках некоторых лингвистических описаний. — ВЯ, 1963, № 1, стр. 113—123.

<sup>12</sup> См.: О. Г. К а р п и н с к а я. Методы типологического описания славянских родовых систем. «Лингвистические исследования по общей и славянской типологии». М., 1966, стр. 105—112.

3. Оканчивается ли И, ед. ч. на *-o* или не *-o* <sup>13</sup>?
  4. Оканчивается ли И, ед. ч. на *-a* или не *-a*?
  5. Оканчивается ли И, ед. ч. на *n*, *r* или остальные согласные?
  6. Присутствует ли в окончании Т, ед. ч. губной носовой звук?
  7. Оканчивается ли Р, мн. ч. на *ŕ*?
  8. Совпадает ли форма В, ед. ч. с формой Р, ед. ч.?
- В словацком языке задаются следующие вопросы:

1. Оканчивается ли И, ед. ч. на *-a* или не *-a*?

Вопросы 2—4 совпадают с соответствующими вопросами в чешском языке.

5. Маркирован ли согласный, на который оканчивается слово в И, ед. ч. по сочетанию признаков «согласность + некомпактность + непалатальность + неяркость» (т. е. оканчивается ли слово на один из следующих согласных: *p, b, m, n, r, d, l, s, z, v, f*)?

6. Оканчивается ли слово в И, ед. ч. на *p, m, v, ŕ* или на другие согласные?

Вопрос 7 совпадает с вопросом 6 для чешского языка, а вопрос 8 с соответствующим вопросом в чешском языке.

9. совпадает ли форма В, мн. ч. с формой Р, мн. ч.?

Схема для чешского языка может быть изложена в виде следующей системы правил:

1. Возьми произвольное существительное *x*. Проверь, оканчивается ли И, ед. ч. на *-i*. Если да, переходи к правилу 8, если нет — к правилу 2.

2. Проверь, оканчивается ли И, ед. ч. на *-e*. Если да, переходи к правилу 6, если нет — к правилу 3.

3. Проверь, оканчивается ли И, ед. ч. на *-o*. Если да, *x* среднего рода, если нет, переходи к правилу 4.

4. Проверь, оканчивается ли И, ед. ч. на *-a*. Если да, переходи к правилу 7, если нет — к правилу 5.

5. Проверь, оканчивается ли И, ед. ч. на *n* или *r*. Если да, переходи к правилу 6\*, если нет — к правилу 8\*.

6. Проверь, присутствует ли губной носовой в форме Т, ед. ч. Если да, переходи к правилу 7, если нет, *x* женского рода.

6\*. Проверь, присутствует ли губной носовой в форме Т, ед. ч. Если да, переходи к правилу 8\*, если нет — *x* женского рода.

7. Проверь, оканчивается ли форма Р, мн. ч. на *-ŕ*. Если да, *x* мужского одушевленного рода, если нет, *x* среднего рода.

7\*. Проверь, оканчивается ли форма Р, мн. ч. на *-ŕ*. Если да, *x* мужского одушевленного рода, если нет *x* женского рода.

8. Проверь, совпадает ли форма В, ед. ч. с формой Р, ед. ч. Если да, то *x* мужского одушевленного рода, если нет, *x* среднего рода.

8\*. Проверь, совпадает ли форма В, ед. ч. с формой Р, ед. ч. Если да, то *x* мужского одушевленного рода, если нет, *x* мужского неодушевленного рода.

Схема для словацкого языка может быть изложена в виде следующей системы правил:

1. Возьми произвольное существительное *x*. Если И, ед. ч. оканчивается на *-a*, то *x* среднего рода. Если нет, переходи к правилу 2.

2. Проверь, оканчивается ли И, ед. ч. на *-e*. Если да, то *x* среднего рода. Если нет, переходи к правилу 3.

3. Проверь, оканчивается ли И, ед. ч. на *-o*. Если да, то *x* среднего рода, если нет, переходи к правилу 4.

4. Проверь, оканчивается ли И, ед. ч. на *-a*. Если да, переходи к правилу 7, если нет — к правилу 5.

5. Проверь, оканчивается ли *x* в И, ед. ч. на один из следующих согласных: *p, b, m, n, r, d, l, t, s, z, v, f*. Если да, переходи к правилу 6, если нет — к правилу 7.

<sup>13</sup> Для простоты общеславянская двузначность окончаний на *-o* и *-e* (имена мужского одушевленного рода в чешском языке и имена лично-мужского в словацком с уменьшительными суффиксами) не учитывается.



6. Проверь, оканчивается ли  $x$  в И, ед. ч. на один из согласных  $p, m, v, r$ . Если да, переходи к правилу 7\*, если нет — к правилу 8.

7. Проверь, присутствует ли губной носовой звук в окончании Т, ед. ч. Если да, переходи к правилу 9, если нет,  $x$  женского рода.

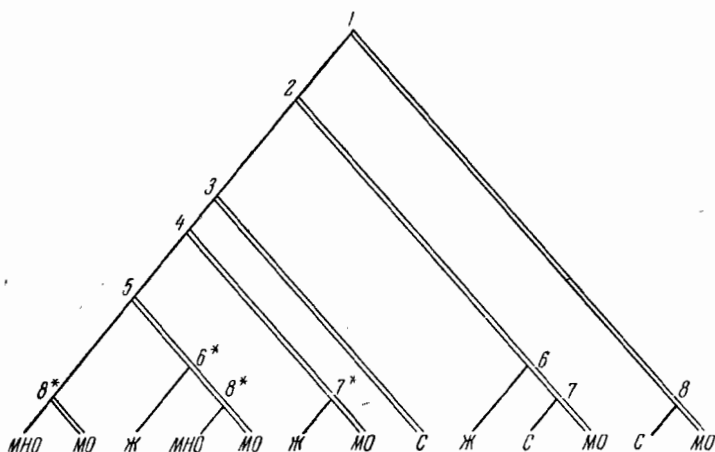
7\*. Проверь, присутствует ли губной носовой в окончании Т, ед. ч. Если да, переходи к правилу 8, если нет,  $x$  женского рода.

8. Проверь, совпадает ли форма В, ед. ч. с формой Р, ед. ч. Если да, переходи к правилу 9\*, если нет —  $x$  мужского неодушевленного рода.

9. Проверь, совпадает ли форма В, мн. ч. с формой Р, мн. ч. Если да,  $x$  лично-мужского рода, если нет,  $x$  среднего рода.

9\*. Проверь, совпадает ли форма В, мн. ч. с формой Р, мн. ч. Если да, то  $x$  лично-мужского рода, если нет,  $x$  нелично-мужского одушевленного (животного) рода.

Порядок, в котором должны задаваться вопросы, и окончательные результаты могут быть представлены следующими схемами (1-я для чешского языка, 2-я — для словацкого), в которых каждый раз справа (отмечено двумя линиями) указываются подмножества, обладающие проверяемым свойством, а слева (отмечено одной линией) подмножества, не обладающие таким свойством.

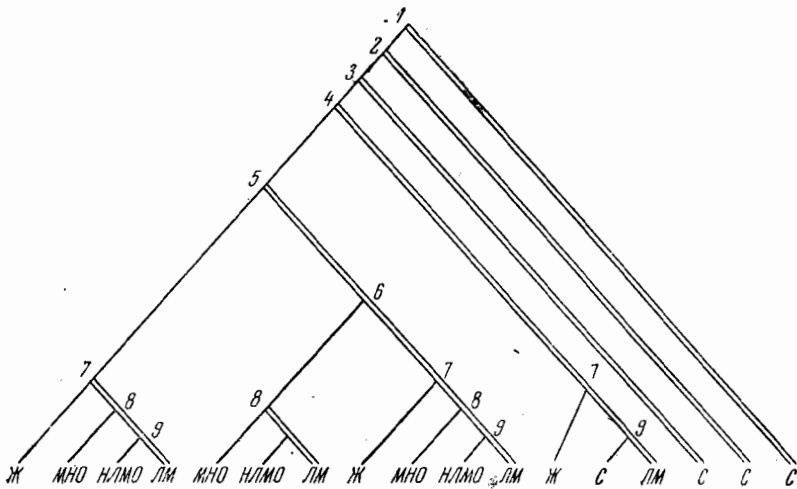


1. Чешский язык

Легко видеть, что мы лишь несколько переформулировали правила, даваемые в грамматиках соответствующих языков. При этом мы получили систему предписаний такого вида, который часто встречается в работах, посвященных лингвистическим приложениям. Мы в сущности построили простейший пример алгоритма. Понятие алгоритма можно определить следующим образом:

- 0\*. 2. 29. А л г о р и т м о м называется конечная система предписаний, а) носящих формальный характер, т. е. формулируемых однозначно, б) расчлененных на ряд элементарных действий, в) применимых к достаточно большой совокупности объектов (массовость), г) приводящих в конечное число шагов к определенному результату (результативность), д) применяемых в строго фиксированном порядке (детерминированность).

Заметим, что определение алгоритма дается под звездочкой. Более точные определения, вернее уточнения понятия алгоритма нас в данной работе



## 2. Словацкий язык

не интересуют, так как нам важно не общее понятие алгоритма, а конкретные алгоритмы.

Пусть теперь имеется некоторое множество элементов  $M_1$ , которое является подмножеством другого множества  $M_2$ .

0. 2. 30. Множество  $M_1$  называется разрешимым, если существует алгоритм, который для любого элемента из  $M_2$  определяет, принадлежит ли этот элемент к  $M_1$ .

Так, если взять в качестве  $M_1$  множество существительных мужского рода чешского языка, а в качестве объемлющего множества  $M_2$  все множество существительных этого языка, то это множество  $M_1$  разрешимо. До сих пор неясно, будут ли лингвисты всегда иметь дело лишь с разрешимыми множествами или же появятся и множества неразрешимые.

Пока по этому вопросу можно высказать лишь предварительные соображения.

1. В большинстве лингвистических задач, связанных с идентификацией каких-то объектов, мы имеем дело с конечными множествами. Так, конечным и даже легко обозримым является множество фонем любого языка. Значительно большим, но все же еще обозримым, а тем более конечным является и множество словоформ любого языка. Для любого конечного множества существует алгоритм решения любой строго формулируемой задачи, например, распознавания какого-либо свойства, ибо на конечном множестве задача в принципе может быть решена методом перебора, т. е. может быть составлен список всех объектов, обладающих нужным свойством, и установлено, находится ли искомым объект в данном списке. Правда, в случаях, когда множество объектов велико (например, когда мы имеем перед собой множество словоформ), практически списком пользоваться невозможно и поэтому целесообразно искать более экономичные алгоритмы, чем метод перебора, хотя полностью этот метод и не может быть исключен (практически любое правило грамматики сопровождается списком «исключений», причем если этот список полон, то пользование им совместно с правилом дает однозначный результат)<sup>14</sup>. В принципе, однако, конечное множество разрешимо.

<sup>14</sup> Ф. Хаусхолдер. Списки в грамматике. «Математическая логика и ее применения». Перев. с англ. под ред. А. И. Мальцева. М., 1965, стр. 293—305.

2. В случаях, когда мы имеем дело с бесконечными множествами, например, с множествами предложений или текстов на данном языке, множества, обладающие некоторым специальным свойством, на разрешимость относительно объемлющей совокупности всех предложений или текстов не проверялись еще, так как для этого нужно прежде всего зафиксировать эти свойства в модели. Соответствующие свойства как правило (ср. однако § 17 о свойстве грамматической правильности) распознаются носителями языка. Поэтому пока большинство исследователей как будто склоняются к гипотезе о том, что человек владеет общим алгоритмом проверки таких свойств<sup>15</sup>.

Этот вопрос еще требует изучения. Строя модели, мы будем во всяком случае стремиться к тому, чтобы формулируемые в них свойства соответствовали разрешимым множествам.

## § 11. ПОРОЖДАЮЩИЙ ПРОЦЕСС И ОПЕРАЦИОННЫЕ ОПРЕДЕЛЕНИЯ

Близко к понятию алгоритма понятие порождающего процесса (или исчисления), который можно определить следующим образом:

**0. 2. 31.** Порождающим процессом называется конечная система предписаний:

- а) носящих операционный характер, т. е. формулируемых однозначно и применяемых к конечному исходному множеству объектов;
- б) расчлененных на ряд элементарных действий;
- в) приводящих к возникновению некоторых новых объектов.

Отличие порождающего процесса от алгоритма модального характера: правила первого предписывают некоторую операцию, а правила второго лишь разрешают провести такую операцию. Поэтому в порождающем процессе порядок применения правил (в отличие от алгоритма) не всегда фиксирован. Поэтому же результатом его действия является не один объект (например, ответ на поставленный вопрос), а конечное или даже бесконечное множество объектов (каждый из которых реализует некоторую возможность). В этой связи интерес приобретает следующее определение.

**0. 2. 32.** Множество  $M$  называется перечислимым, если существует процесс, порождающий все элементы этого множества.

**Т. 2. 2.** Всякое разрешимое множество перечислимо, но обратное неверно: существуют перечислимые множества, не являющиеся разрешимыми.

Примеры лингвистически значимых порождающих процессов мы рассмотрим в § 20 и далее более подробно в главе VII. Здесь же приведем более простой пример. Пусть нам заданы следующие объекты, сгруппированные по классам.

$A_1$	$A_2$	$A_3$	$A_4$	$A_5$
<i>один</i>	<i>пять</i>	<i>десять</i>	<i>двадцать</i>	<i>сто</i>
<i>два</i>	<i>шесть</i>		<i>тридцать</i>	<i>двести</i>
<i>три</i>	<i>семь</i>		<i>сорок</i>	<i>триста</i>
<i>четыре</i>	<i>восемь</i>		<i>девяноста</i>	<i>четыреста</i>
	<i>девять</i>			

Исходя из этих первоначальных объектов (число их можно, разумеется, еще уменьшить), мы можем произвести названия всех числительных от 1 до 999<sup>16</sup>.

<sup>15</sup> См.: Х. П у т н а м. Некоторые спорные вопросы теории грамматики. Перев. с англ. «Новое в лингвистике», вып. IV. М., 1965, стр. 66—96.

<sup>16</sup> Наряду с *девяноста* мы произведем и более регулярную, хотя с современной точки зрения и неправильную форму *девятыдесять* (ср. параллельное упоминание обеих форм в словаре Срезневского).

1) Возьми любой из объектов группы  $A_2$ , припиши к нему без пробела справа *десять* и помести получившийся объект в группу  $A_4$ .

2) Возьми любой из объектов группы  $A_1$  или  $A_2$  и если он оканчивается на *-ь* или *-е*, то отбрось эти буквы, а если он оканчивается на *-а*, то замени *-а* на *-е*; припиши к получившемуся объекту справа без пробела *-надцать* и помести получившийся объект в  $A_3$ .

3) Возьми любой из объектов группы  $A_2$ , припиши к нему справа без пробела *сот* и помести получившийся объект в группу  $A_5$ .

4) Возьми произвольный объект группы  $A_1$  или  $A_2$  и остановись или перейди к 5.

5) Припиши слева к объекту группы  $A_1$  или  $A_2$  какой-нибудь объект группы  $A_4$  и остановись или перейди к 6.

6) Возьми какой-нибудь объект, получившийся по правилу 4 или 5 и припиши к нему слева объект группы  $A_5$ .

Например, по этим правилам число *семьсот пятьдесят три* получается так: сначала по правилу 1 образуется число *пятьдесят* и по правилу 3 число *семьсот*. Затем из группы  $A_1$  берется число *три*, к нему по правилу 5 приписывается слева *пятьдесят*, а затем к получившейся последовательности *пятьдесят три* по правилу 6 приписывается *семьсот*.

Эти правила можно было бы дополнить так, чтобы получалось произвольное количественное числительное русского языка.

Заметим, что порождающий процесс может быть истолкован как способ определения получающихся при этом объектов. В самом деле можно дать следующее определение числительного, состоящее из двух частей:

а) слова *один, два, три, четыре, пять, шесть, семь, восемь, девять, десять, двадцать, тридцать, сорок, девяносто, сто* — суть числительные русского языка;

б) если некоторое слово образовано по правилам 1—6 из числительных русского языка, то оно также есть числительное русского языка.

Такого рода определения называются *операционными*, или *индуктивными*, и мы не раз еще столкнемся с такими определениями.

Особенность их состоит в том, что в первой части называется некоторая совокупность объектов, подпадающих под определение (эта часть называется *базисом* определения), а во второй объекты определяются при помощи получаемых по определению объектов. Но здесь нет никакого порочного круга, ибо постепенно расширяется совокупность объектов, подпадающих под определение.

Операционные определения похожи на методические приемы обучения языку: сначала даются примеры, а потом эти примеры обобщаются (так *Ваня играет* есть предложение и любая последовательность слов, полученная заменой слова *Ваня* на имя другого человека, также есть предложение).

В лингвистической теории до сих пор однако стремились к определениям другого рода, а именно «видо-родовым», т. е. таким, в которых указывается объемлющая совокупность объектов («род») и некоторое подмножество, обладающее в нем особым свойством («вид»). В качестве примера можно привести определение морфемы: «Морфема есть минимальная значимая единица языка». Здесь предполагается, что имеется объемлющая совокупность, а именно множество значимых единиц языка и в ней подмножество единиц, обладающих свойством минимальности. Однако такие определения строить довольно трудно (например, для «предложения» до сих пор нет приемлемого видо-родового определения), а главное, такие определения не дают в целом ряде случаев критериев для четкого выделения объекта (ведаром в школе не пользуются «научным» определением подлежащего, дополнения и т. п., а предлагают пользоваться методом «вопросов»). Между тем, как высечено в современной логике, четкое (в частности, построенное на принципе объемности — см. § 6) видородовое определение определяет некоторое разрешимое множество, в то время как индуктивное определение определяет некоторое перечислимое

множество<sup>17</sup>. Это означает, что корректное с формальной точки зрения видородовое определение может быть перестроено как индуктивное, в то время как обратное, вообще говоря, неверно, т. е. существуют такие индуктивные определения, которые не могут быть перестроены в видородовые. Есть основания предполагать, что таково, например, понятие «предложения», которое, по-видимому, можно определить лишь индуктивно<sup>18</sup>.

Во всяком случае мы будем в дальнейшем широко пользоваться не только видородовыми определениями, но также и индуктивными.

## § 12. ПОНЯТИЯ РАССТОЯНИЯ И ПРИЗНАКОВОГО ПРОСТРАНСТВА

0. 2. 33. Некоторое множество  $M$  называется метрическим пространством, если любой паре его элементов  $x$  и  $y$  поставлено в соответствие некоторое число  $\rho(x, y)$ , обладающее следующими свойствами:

- 1)  $\rho(x, y) = \rho(y, x)$  — симметричность
- 2)  $\rho(x, y) \geq 0$  — неотрицательность
- 3)  $\rho(x, y) = 0$  тогда и только тогда, когда  $x = y$  — невырожденность
- 4) для любого  $z$   $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$  — аксиома треугольника

Число  $\rho$  называется расстоянием между  $x$  и  $y$ . Различные примеры расстояний описываются, например, в работе Ю. А. Шрейдера<sup>19</sup>. Наиболее важно с лингвистической точки зрения описываемое им пространство сообщений и определяемая для этого пространства функция расстояния, ибо везде, где в лингвистике возникает необходимость измерять меру близости двух объектов, можно воспользоваться именно такой или близкой к ней функцией расстояния<sup>20</sup>.

Каждое сообщение может быть представлено как упорядоченная последовательность элементов некоторого фиксированного множества (которое в этом случае принято называть алфавитом). Пусть все места в сообщениях перенумерованы (в каждом начиная с единицы).

0. 2. 34. Расстоянием между двумя сообщениями называется число мест, на которых стоят разные элементы.

Легко видеть, почему именно такая функция расстояния удобна для большинства лингвистических задач. Мы уже говорили, что в лингвистике мы часто имеем дело с объектами, характеризуемыми некоторой совокупностью признаков, причем каждый из таких признаков принимает несколько значений (в простейшем случае ровно два значения). Записав все значения признаков в определенном алфавите (например, для простейшего случая, когда каждый признак принимает ровно два значения: можно ограничиться алфавитом, состоящим из нуля — отсутствие признака и единицы — наличие признака) и зафиксировав порядок признаков, мы можем поставить в соответствие каждому объекту определенное сообщение (информацию о соответственных значениях каждого признака), а тем самым и измерять расстояние между двумя произвольными объектами. Поэтому совокупность объектов с заданными признаками и их значениями можно назвать «признаковым пространством».

<sup>17</sup> См. статью В. А. Успенского «Алгоритмы» в «Философской энциклопедии».

<sup>18</sup> См.: И. Бар-Хиллел. О рекурсивных определениях в эмпирических науках. «Математическая лингвистика» (сб. переводов). М., 1964, стр. 101—107.

<sup>19</sup> Ю. А. Шрейдер. Что такое расстояние. М., 1963.

<sup>20</sup> Хотя, разумеется, существуют и другие средства. См., например: Ю. И. Левин. Об описании системы лингвистических объектов, обладающих общими свойствами. — ВЯ, № 4, 1964.

В качестве примера рассмотрим расстояние между русскими согласными  $p, b, t, d, k, g$ .

Для них мы имеем следующую таблицу признаков:

	$p$	$b$	$t$	$d$	$k$	$g$
Компактность	0	0	0	0	1	1
Периферийность	1	1	0	0	1	1
Звонкость	0	1	0	1	0	1

Каждый столбец представляет собой с формальной точки зрения некоторое сообщение. Расстояния между ними можно изобразить так:

	$p$	$b$	$t$	$d$	$k$	$g$
$p$	0	1	1	2	1	2
$b$	1	0	2	1	2	1
$t$	1	2	0	1	2	3
$d$	2	1	1	0	3	2
$k$	1	2	2	3	0	1
$g$	2	1	3	2	1	0

Ясно, что приведенный нами пример тривиален и что для подобных сопоставлений нет нужды вводить понятие расстояния. Однако уже для фонологии понятие расстояния и вводимые им аналогии (например, с пространством сообщений) могут привести к интересным результатам<sup>21</sup>. Мы ограничимся здесь всего одним простым примером<sup>22</sup>. В приведенной нами таблице расстояние между некоторыми звуками, например, между  $p$  и  $b$  или  $t$  и  $d$  равно 1.

<sup>21</sup> См.: Вяч. Вс. Иванов.  $n$ -мерное пространство языка. «Бюллетень объединения по проблемам машинного перевода». М., 1957, № 5, стр. 51—52; а также: E. Colin Cherry. Roman Jakobson's «Distinctive Features» as the normal coordinates of a language. «For Roman Jakobson». The Hague, 1956.

<sup>22</sup> Этот пример более подробно см.: И. И. Ревзин. К логическому обоснованию теории фонологических признаков. — ВЯ, 1964, № 5, в связи с обсуждением теории, выдвинутой в работе: S. Marcus. Un model matematic al fonemului. «Studii și cercetări matematice», vol. XIV, № 3, 1963; см. также: S. Marcus. Lingvistică matematică. București, 1963, стр. 56—68.

Между тем в теории передачи сообщений исходят из оправдываемой предпосылки, что два сообщения, расстояние между которыми равно единице, часто смешиваются, т. е. одно из них может быть воспринято как другое. Возникает вопрос, почему не смешиваются между собой звуки *p* и *b* или *p* и *k*. Ответ состоит в том, что здесь учитываются не все признаки, различающие два звука. Это подтверждается и новейшими экспериментально-фонетическими исследованиями, в которых выяснено, что в большинстве случаев, когда по таблице дифференциальных признаков расстояние минимально, на самом деле противопоставляемые пары различаются не одним, а несколькими признаками<sup>23</sup>.

Таким образом, рассмотрение расстояний между звуками приводит к теоретическим выводам, которые хорошо согласуются с новейшими экспериментально-фонетическими данными.

Понятие расстояния, как мы увидим в главе VI, может быть полезным и при исследовании грамматических категорий.

### § 13. ПРЕДВАРИТЕЛЬНЫЕ ЛИНГВИСТИЧЕСКИЕ ПРИМЕРЫ (НЕКОТОРЫЕ ПОНЯТИЯ ФОНОЛОГИИ)

Фонология является наиболее развитой и логически совершенной областью языкознания, поэтому на ее примере лучше всего иллюстрируются введенные понятия (часть примеров мы уже привели раньше — здесь же мы дадим более связное изложение фонологических понятий).

Пусть нам задано множество звуков (мы уже договорились в § 1, что, говоря о звуках, мы будем иметь в виду звуки-типы; здесь же мы уточним, что нам звуки речи даны как звуки языка в понимании П. С. Кузнецова<sup>24</sup>, т. е., что разделение всего множества индивидуальных звуков речи на звуки-типы представляет собой разбиение).

Пусть на множестве звуков-типов (звуков языка) задано *n* разных отношений линейного порядка  $O_i$ . Тогда исходное множество звуков языка можно представить как ряд линейно упорядоченных последовательностей, некоторые из которых можно интерпретировать как фонетические слова, в то время как другие суть последовательности звуков, невозможные в данном языке.

О. 2. 35. Пусть задано множество десигнатов (см. § 1) слов. Будем считать, что определено некоторое отображение, сопоставляющее некоторым последовательностям один и только один десигнат (для простоты мы отвлекаемся от случаев омонимии). Тогда выделится некоторое подмножество последовательностей, являющихся областью определения данного функционального отображения и каждая последовательность из этого подмножества называется словоформой, или, как мы будем говорить в дальнейшем, — словом.

<sup>23</sup> См.: Л. В. Бондарко, Л. Р. Зиндер. О некоторых дифференциальных признаках русских согласных фонем. — ВЯ, 1966, № 1, стр. 10—14.

<sup>24</sup> См.: П. С. Кузнецов. Об основных положениях фонологии. — ВЯ, 1959, № 2, стр. 30—31.

Рассмотрим теперь некоторые другие отношения, определяемые на множестве звуков.

**О. 2. 36.** Два звука  $x$  и  $y$  находятся в отношении к о м м у т а ц и и  $K$ , если существуют такие слова  $AxB$  и  $AyB$ , которые поставлены в соответственные разным десигнатам.

Исследуем отношение  $K$ . Оно, очевидно, антирефлексивно и симметрично. Легко построить пример, показывающий, что это отношение нетранзитивно. Нетранзитивно<sup>25</sup> и дополнительное к нему отношение  $\bar{K}$ . Это новое отношение интересно, однако, со следующей точки зрения; его можно интерпретировать как критерий отнесения двух звуков к одной фонеме, т. е. считать, что два звука относятся к одной фонеме, если не существует двух разных слов, различающихся только этими звуками.

Но разбивает ли отношение  $\bar{K}$  все множество звуков на непересекающиеся классы? Очевидно, не разбивает, ибо оно нетранзитивно. Правда, для тех теорий фонемы, которые исходят из предпосылки, что звуки, относящиеся к разным фонемам, не должны совпадать, можно либо просто постулировать транзитивность коммутации, либо попытаться формальным путем построить на базе данного отношения новое отношение, которое является отношением эквивалентности, используя операцию транзитивного замыкания отношения  $\bar{K}$ <sup>26</sup>.

Однако, как следует из цитированной работы В. А. Успенского, который в свою очередь опирается на примеры Глиссона<sup>27</sup>, подобное определение фонемы приведет к результатам, противоречащим принципу смыслоразличения, который лежит в основе понятия фонемы.

Это легко видно и из следующего простого примера (лишь немного идеализирующего ситуацию в таких языках, как тюркские). Пусть в языке два гласных и два согласных не могут следовать один за другим. Тогда любой гласный находится в отношении  $\bar{K}$  к любому согласному и наоборот. Но тогда при транзитивации все гласные и все согласные объединятся в один класс, что показывает полную бессмысленность такого подхода.

Таким образом, транзитивация в этом случае ничего не дает. Но мы не знаем каких-либо других способов превращения данного отношения в отношение эквивалентности, что затрудняет определение принадлежности к одной фонеме через это отношение.

<sup>25</sup> Ср. исследование отношения «не иметь минимальных пар» в работе: В. А. Успенский. Одна модель для понятия фонемы. — ВЯ, 1964, № 6, стр. 47.

<sup>26</sup> Несколько отличная процедура приводится в работе: R. A b e r n a t h y. The problem of linguistic equivalence. — PSAM, vol. 12. «Structure of language and its mathematical aspects». Providence, 1961, стр. 95—98.

<sup>27</sup> Г. Глиссон. Введение в дескриптивную лингвистику. Перев. с англ. М., 1959, стр. 253.



0. 2. 37. Два звука  $x$  и  $y$  находятся в отношении дополнительного распределения, если не существует таких последовательностей  $AxB$  и  $AyB$ , которые одновременно являются словами данного языка.

Это отношение антирефлексивно, симметрично, но также не транзитивно. Если два звука находятся в отношении дополнительного распределения, то они находятся в отношении  $\bar{K}$ , т. е. дополнительном к отношению коммутации. Обратное, разумеется, не имеет места.

0. 2. 38. Два звука  $x$  и  $y$  находятся в отношении свободного варьирования, если для любых последовательностей  $A$  и  $B$  слова  $AxB$  и  $AyB$ , либо оба не имеют значения, либо имеют одинаковое значение.

Отношение свободного варьирования есть отношение типа эквивалентности. При этом если два звука находятся в отношении свободного варьирования, то они находятся в отношении  $\bar{K}$ .

Отношение свободного варьирования обычно служит для объединения похожих звуков в один абстрактный звук-тип, однако для получения аналога понятия фонемы его недостаточно.

Итак, определение разбиения всего множества звуков на классы, соответствующие фонемам, через некоторое отношение, связанное со смыслоразличительной ролью фонем, оказывается достаточно трудным. Возможен, однако, обратный подход к определению фонемы. Предположим, что нам дано разбиение всего множества звуков на фонемы, и посмотрим, какими свойствами должно обладать это отношение.

Однако прежде всего введем еще некоторые дополнительные понятия. Поскольку каждой последовательности звуков соответствует фонемная цепочка, то вполне естественно определить отношение коммутации не только для последовательностей звуков, но и для последовательностей фонем, ибо каждому слову в звуковом представлении соответствует слово в фонемном представлении. Итак, пусть нам дано отображение  $F$ , ставящее в соответствие каждому звуку его фонему, и, кроме того, на множестве звуков определено свое отношение коммутации, а на множестве фонем — свое.

Естественно при этом потребовать, чтобы отображение  $F$  было гомоморфным относительно отношения коммутации.

Таким образом, процедура определения фонемы должна сводиться: а) к формулировке некоторой гипотезы о возможном разбиении на фонемы и б) к проверке этой гипотезы<sup>28</sup>.

Практически гипотеза о возможном разбиении на фонемы формируется на основе тех фонетических признаков, которые

---

<sup>28</sup> Подобный подход был впервые четко сформулирован у нас С. К. Шамяном, но автор пришел к нему совсем другим путем, а именно через попытку осмыслить цитируемую выше статью В. А. Успенского.

сопоставляются каждому звуку и о которых мы говорили в § 10. А именно, в один класс объединяются звуки, у которых большинство признаков совпадает, при этом несовпадающие признаки временно (до подтверждения или отклонения гипотезы) считаются неважными.

Указанное обстоятельство, по-видимому, повлияло и на понимание фонемы в большинстве современных концепций не как класса звуков, а как совокупности или «пучка» различительных признаков.

Последнее важно для нас и потому, что позволяет изложить теорию оппозиций Трубецкого в терминах теории множеств<sup>29</sup>. Это тем более удобно, что теория оппозиций применяется не только в фонологии, но и в грамматике и вообще может быть использована для произвольной системы лингвистических объектов, каждому из которых поставлена в соответствие некоторая совокупность свойств. Мы будем обозначать здесь объекты строчными латинскими буквами  $a$  и  $b$ , а поставленные им в соответствие множества свойств — заглавными буквами  $A$  и  $B$ .

0. 2. 39. Если  $A$  и  $B$  два непустых множества свойств и имеет место  $A \subseteq B$ , то говорим, что оппозиция между  $a$  и  $b$  является *п р и в а т и в н о й* в пользу  $b$ .

0. 2. 40. Б а з о й оппозиции  $a$  и  $b$  называется множество  $A \cap B$ .

0. 2. 41. Если  $A$  и  $B$  два непустых множества свойств, причем не имеет место ни  $A \subseteq B$ , ни  $B \subseteq A$ , то оппозиция между  $a$  и  $b$  называется *э к в и п о л е н т н о й*, если ее база непуста, и *д и з ъ ю н к т и в н о й* в противном случае.

0. 2. 42. Если имеются две оппозиции с одинаковой базой, то мы говорим, что имеет место *м н о г о с т о р о н н я я* оппозиция.

Чрезвычайно важным видом оппозиций, как известно, являются пропорциональные оппозиции (в отличие от изолированных).

Это понятие можно ввести следующим образом<sup>30</sup>.

0. 2. 43. Пусть имеется две пары оппозиций: оппозиция  $a_1$  и  $b_1$  и оппозиция  $a_2$  и  $b_2$ . Возьмем те элементы  $A_1$  и  $B_1$ , которые не входят в базу этой оппозиции, т. е. в  $A_1 \cap B_1$ . Если получившееся множество элементов совпадает с множеством тех элементов  $A_2$  и  $B_2$ , которые не входят в  $A_2 \cap B_2$ , то мы будем говорить, что оппозиция  $(a_1, b_1)$  *п р о п о р ц и о н а л ь н а* оппозиции  $(a_2, b_2)$ .

<sup>29</sup> Эта задача была выполнена в цитированных выше работах Кантино и Маркуса (см. сноску 1, стр. 33).

<sup>30</sup> Наш способ формализации этой оппозиции лишь по форме отличается от предложенного С. Маркусом, см.: С. М а р к у с. Логический аспект лингвистических оппозиций. «Проблемы структурной лингвистики. 1963». М., 1963, стр. 54; см. также: S. M a r c u s. *Lingvistică matematică*, стр. 15.

0. 2. 44. Оппозиция  $(a_1, b_1)$  называется *изолированной*, если не существует никакой другой оппозиции, пропорциональной ей.

Возьмем, например, фонемы  $p$  и  $b$ . Фонеме  $p$  соответствует множество признаков: (некомпактность, недиезность, периферийность, незвонкость, неяркость), фонеме  $b$  соответствует множество признаков (некомпактность, недиезность, периферийность, звонкость, неяркость). База оппозиции есть множество (некомпактность, недиезность, периферийность, неяркость). Признаки, не входящие в базу, суть звонкость и незвонкость. Если мы сравним совокупность признаков, общих для  $t$  и  $d$ , для  $k$  и  $g$ , то получим те же признаки. Поэтому оппозиция  $(p, b)$  пропорциональна оппозиции  $(t, d)$  и пропорциональна оппозиции  $k$  и  $g$ . Примером изолированной оппозиции в русском языке может служить оппозиция  $(t, c)$ , поскольку фонеме  $c$  соответствует совокупность признаков (некомпактность, диезность, непериферийность, глухость, яркость), а фонеме  $t$  совокупность признаков (некомпактность, недиезность, непериферийность, неяркость) и совокупность признаков, не входящих в базу оппозиции (диезность, яркость, недиезность, неяркость), не получается ни для какой другой оппозиции в русском языке.

Других примеров из фонологии мы приводить не будем, поскольку теория оппозиций хорошо знакома каждому лингвисту.

Мы хотим однако обратить внимание на другое, а именно на ценность разграничения пропорциональных и изолированных оппозиций для грамматики. Как подчеркнул Кантино, это разграничение лежит в основе грамматики, а именно пропорциональные оппозиции относятся к грамматике, а изолированные к лексике<sup>31</sup>.

Следуя цитированной работе Кантино, оппозиции значимых единиц можно формализовать следующим образом.

К числу свойств каждого слова следует отнести:

- а) то, что оно имеет определенный десигнат,
- б) то, что в нем выступают определенные формальные показатели, например, окончания,
- в) то, что ему сопоставлены определенные грамматические категории.

Например свойства слова *волку* можно записать следующим образом: {десигнат *lupus*, окончание *-у*, Д, М, ед. ч.}, а свойства слова *столами* следующим образом {десигнат *mensa*, окончание *-ами*, Тв, М. мн. ч.}. Аналогично можно выписать свойства для слов *столу* и *волками*. Тогда легко видеть, что совокупность свойств, не входящих в базу, для обеих оппозиций *столу—сто-*

<sup>31</sup> J. C a n t i n e a u. Les oppositions significatives. «Cahiers Ferdinand de Saussure», 10, 1952, стр. 27.

лами и волку—волками одна и та же, а именно {окончание -у, Д, ед. ч., окончание -ами, Тв; мн. ч.}. Поэтому эти оппозиции пропорциональны друг друга.

Изолированной при таком подходе является, например, оппозиция *есть*—*быть*. Но это явное исключение, вообще же все парадигмы построены по принципу пропорциональности. Это замечание важно потому, что в дальнейшем будет использован близкий к этой идее принцип однородности, который будет одним из основных инструментов моделирования грамматики.

## ОСНОВНЫЕ ЛИНГВИСТИЧЕСКИЕ ПОНЯТИЯ И ОПЕРАЦИИ

### § 14. СЛОВО И ПАРАДИГМА

В дальнейшем мы будем часто пользоваться термином «слово». При этом «слово» можно было бы просто определять как последовательность звуков, поставленную в соответствие дезигнату (см. определение 2. 35). Однако в теории моделей целесообразнее использовать это понятие как неопределяемое, заданное извне. Иначе говоря, мы считаем, что нам задано некоторое исходное множество слов  $V = \{x\}$ , которое называется словарем и элементы которого называются словами. Интерпретируется «слово» как словоформа, обычно синтетическая («от пробела до пробела»), но в некоторых случаях, где это специально будет оговорено, и аналитическая (сочетание специального вида служебных слов с «полнозначными»). Таким образом, *стол* и *стола*, *гуляю* и *буду гулять* интерпретируются как разные слова, а последнее может пониматься двояко как два слова или как одно. Способы проведения границ между словами, т. е. сегментации текста, не рассматриваются в данной работе<sup>1</sup>.

Подобное употребление термина «слово» делает необходимым ввести аналог и для такого употребления термина «слово», когда оно рассматривается как определенная система словоформ. Для этого мы введем еще один неопределяемый термин, а именно «парадигма». Можно было бы ввести это понятие следующим образом. Большинство слов в нашем смысле (например, *смола*, *гулять*) обозначают какие-то объекты или явления внеязыковой действительности, это так называемые *полнозначные* слова. Имеется определенный закон соответствия между словами и объектами, иначе говоря, имеется некоторое отображение  $\psi$ , сопоставляющее каждому слову совокупность объектов или явлений действительности, или значение этого слова (в смысле Фреге). Можно для простоты считать, что каждому слову ста-

<sup>1</sup> Этот вопрос рассматривался автором в «Моделях языка» (стр. 54—56). С тех пор появились более удачные модели сегментации, например, модель, предложенная П. С. Кузнецовым (П. С. Кузнецов. Опыт формального определения слова. — ВЯ, 1964, № 5).

вится в соответствие ровно одно значение. Тогда можно бы было сказать, что в одну парадигму попадают все слова, поставленные в соответствие одному объекту или явлению действительности (например, *стол, стола, столу, столом* и т. п.). Иначе говоря, каждая парадигма есть полный прообраз некоторого объекта или явления действительности при отображении  $\psi$ .

Такое введение термина «парадигма» пролило бы свет на его связь с ввелингвистическими факторами, однако оно связано с существенными трудностями. Во-первых, отображение  $\psi$ , о котором шла речь выше, не может быть определено на множестве всех слов, ибо имеется целый класс неполнозначных слов, не обозначающих никаких объектов или явлений действительности. Эту трудность можно было бы, разумеется, преодолеть, введя особое значение для каждого из неполнозначных слов, что, собственно говоря, и предполагалось в определении 2. 35. Во-вторых, при таком понимании термина «парадигма», разбиение на парадигмы однозначно определяется данным внешним миром. Между тем решение вопроса о том, являются ли два слова формами друг друга или нет во многом зависит от внутриязыковых соображений и поэтому целесообразно допустить определенную степень свободы в определении объема парадигм.

Поэтому более удобно считать, что извне задано разбиение на парадигмы, соответствующее тому принципу, что в одну парадигму попадают слова с общей основой. Ниже мы остановимся на некоторых принципах выделения более мелких парадигм и, в частности, обсудим вопрос о том, следует ли объединять в одну парадигму формы единственного и множественного числа.

**Примечание.** Ранее автор вместо термина «парадигма» использовал термин «окрестность», который явно неудачен.

## § 15. ГРАММАТИЧЕСКИЕ КАТЕГОРИИ

К числу исходных, неопределяемых понятий теории моделей, вообще говоря, следует отнести и понятие грамматической категории. Под грамматической категорией, мы будем следуя Пешковскому<sup>2</sup> понимать некоторое множество слов, объединяемое какой-то общей формальной чертой. Правда, ниже в главе VI будут рассмотрены процедуры, дающие возможность отнести каждое существительное к определенному роду и определенному падежу (вернее, аналогам рода и падежа). Но не все грамматические категории, важные для дальнейшего описания, можно опре-

<sup>2</sup> А. М. Пешковский. Русский синтаксис в научном освещении. Изд. 6. М., 1938, стр. 53—61. — Концепция Пешковского осталась, по-видимому, совершенно незамеченной на Западе, ср. обсуждение вопроса о категориях: «*Zeichen und System der Sprache*», Bd. II, Berlin, 1960, стр. 80 и сл.

делить подобными процедурами (например, пока не удалось и неизвестно, удастся ли определить так категорию числа).

Кроме того, мы выбрали такой подход, при котором принадлежность слова к определенным категориям первоначально считается заданной, ибо это даст нам возможность лучше выяснить отношения между категориями и другими понятиями. Знание общих свойств категорий поможет нам в главе VI выводить некоторые категории из более примитивных понятий.

Поскольку мы считаем принадлежность к категориям, т. е. к определенным множествам слов, заданной, нам придется ввести некоторые условия относительно омонимии, а именно омонимы, принадлежащие к разным частям речи, например, слова типа *печь* мы с самого начала будем считать двумя различными словами, снабженными соответствующими индексами, например, *печь*<sub>1</sub> и *печь*<sub>2</sub> (о возможности формального выделения этого класса омонимов без обращения к понятиям части речи и категории см. § 25). Мы будем считать, что для каждого из таких омонимов категории задаются отдельно. Если же слово принадлежит к двум разным падежам (например, слово *стол*, принадлежащее к именительному и винительному падежам), то мы введем для него специальные смешанные категории (например, именительно-винительного падежа). Такой способ задания категорий не отличается от того, который получается при строго морфологическом подходе к грамматическим категориям (например, категории падежа)<sup>3</sup>. Эти условности связаны лишь с первоначальным заданием категорий. В дальнейшем будут введены процедуры, приводящие к более естественному объединению слов в категории.

Принятое нами употребление термина «категория» дает возможность говорить о категории определенного рода, например женского, но не категории рода как таковой, или категории определенного падежа, например дательного, а не категории падежа как таковой.

В наших терминах сочетание «категория рода» бессмысленно, так как если оно что-либо обозначает, то лишь объединение слов, имеющих род, но это то же объединение, что и совокупность слов, имеющих падеж, и тогда «категория рода» совпадает с «категорией падежа», т. е. по принятому нами принципу объемности (см. § 6) свойства «иметь род» и «иметь падеж» были бы тождественными<sup>4</sup> (определяя одно и то же множество слов). Поэтому мы говорим

<sup>3</sup> Ср. типы парадигм в работе: З. М. Волоцкая, Т. Н. Молошная, Т. М. Николаева, И. И. Ревзин, Т. В. Цивьян. Об одном подходе к типологии славянских языков. «Славянское языкознание. Доклады советской делегации. V Международный съезд славистов». М., 1963, стр. 544—547.

<sup>4</sup> Как будет ясно из дальнейшего (глава VI), мы рассматриваем категорию рода у существительных, прилагательных и глаголов как принципиально разные образования.

о «категориях рода» или «категориях падежа», но не о «категориях рода» и «категории падежа».

Чем же в наших терминах является «категория рода вообще» или «категория падежа вообще»? Мы предпочитаем говорить о функции рода или функции падежа и т. п. «Род вообще» или «падеж вообще» в наших терминах есть отображение, определенное на некотором множестве слов (мы увидим дальше, что такие отображения можно определять на всем множестве слов) и ставящее в соответствие каждому слову категорию определенного рода, падежа и т. п.

Существует и другая традиция, в которой слово «категория» употребляется для нашего термина «функция», например «категория числа», «категория падежа» и т. п., в то время как для категории определенного числа или падежа употребляется термин «категориальная форма». Эту традицию, идущую, по-видимому, от А. И. Смирницкого<sup>5</sup>, недавно изложил в четком формализованном виде Б. А. Успенский<sup>6</sup>. При этом категория также понимается как класс элементов (не слов, а служебных элементов), взаимозаменяемых в определенной позиции. Эта концепция, о которой мы еще будем говорить в следующем параграфе, имеет тот недостаток, что она не дает возможности рассматривать как категории те объединения, в которых грамматическая принадлежность определенного слова к классу не обозначается специальным служебным элементом (например, окончанием в славянских языках). Тогда отпадает возможность рассматривать такие категории как категория рода в славянских языках и вообще все те категории, которые Б. Уорф удачно обозначил термином «скрытые категории»<sup>7</sup> (*covert categories*) и которые играют, как показал тот же автор, важную роль в типологическом анализе.

Таким образом, мы считаем заданными ряд отображений, определенных на одном и том же множестве слов:  $\Phi_1, \Phi_2, \dots, \Phi_n$ ; естественно при этом определить новую функцию  $\Pi$ , которая ставит в соответствие каждому слову пересечение всех категорий, в которые входит это слово<sup>8</sup>.

**0. 3. 1.** Пересечение всех категорий, в которые входит данное слово, мы будем называть категориальным пучком<sup>9</sup>.

<sup>5</sup> А. И. Смирницкий. Синтаксис английского языка. М., 1957.

<sup>6</sup> Б. А. Успенский. Структурная типология языков. М., 1965, стр. 79—80.

<sup>7</sup> В. Л. Уорф. Grammatical categories. «Language», vol. 20, № 1, 1945.

<sup>8</sup> Исследование этой и ряда других близких функций см.: С. Я. Фитильов. О построении формальной морфологии в связи с машинным переводом. «Доклады на конференции по обработке информации, машинному переводу и автоматическому чтению текста», вып. 2. М., 1961.

<sup>9</sup> В «Моделях языка» в этом смысле употреблялся термин «элементарная грамматическая категория», имеющий, однако, в большинстве работ по математической лингвистике другое значение.



Мы будем обозначать категориальный пучок слова  $x$  через  $\Pi(x)$  и пользоваться следующей системой записи: если  $\Pi(x) = \kappa_1 \cap \dots \cap \kappa_n$ , где  $\kappa_i$  категории, в которые входит слово  $x$ , то мы будем писать:  $\Pi(x) = (\kappa_1, \dots, \kappa_n)$ . Это делается для того, чтобы не употреблять знак пересечения  $\cap$  в сочетании с грамматическими символами. Так, вместо  $(Ж \cap \text{ед. ч.} \cap И)$  мы будем пользоваться более обычной записью:  $(Ж, \text{ед. ч.}, И)$ .

## § 16. ПОНЯТИЯ ЦЕПОЧКИ И ВХОЖДЕНИЯ

В § 1 мы уже говорили о том, что каждый элемент семиотической системы, в том числе и естественного языка, участвует в двух рядах отношений, а именно в парадигматических и в синтагматических отношениях. Последние представляют собой отношения звуков к другим звукам в слове, отношения слова к другим словам во фразе. Поэтому в теории моделей необходимо понятие, которое обобщает понятие фразы или предложения; а именно понятие произвольной последовательности элементов, вне зависимости от того, является ли она правильной или неправильной. В «Моделях языка» (и в некоторых других работах, посвященных моделированию) для этого понятия употреблялся термин «фраза», однако такое словоупотребление вряд ли войдет в лингвистику. Осознав это обстоятельство<sup>10</sup>, автор предложил употреблять в этом смысле термин «кортеж»<sup>11</sup>, но и этот термин имеет мало шансов на употребление в лингвистике. В советских работах по порождающим грамматикам начал употребляться термин «цепочка» (соответствие английскому string), который мы и будем употреблять, надеясь, что, наконец, удастся достичь единства терминологии лингвистов и математиков, работающих в данной области.

В этой книге термин «фраза» будет употребляться лишь в случаях, когда речь будет идти о последовательностях слов, правильных грамматически или семантически, а в общем случае будет употребляться термин «цепочка».

Цепочка может соответствовать: а) предложению, например, *очень маленький мальчик гуляет*; б) части предложения, например, *я подошел к*, в) бессмысленному ряду слов *я к видеть но уж санями*, г) одному слову, например, *саней* и, кроме того, для

<sup>10</sup> См.: И. И. Ревзин. Некоторые замечания о методах введения математических терминов в лингвистику. «Славянска лингвистична терминология». София, 1962, стр. 61.

<sup>11</sup> По-видимому, под влиянием лекций В. А. Успенского. Термин «кортеж» в том же смысле, но с более широкой сферой употребления используется в кн.: Ю. А. Шиханович. Введение в современную математику. Начальные понятия. М., 1965, стр. 89—118.

общности мы будем считать, что цепочка может быть пустой (пустую цепочку мы будем обозначать символом  $\Lambda$ ).

Итак мы принимаем следующее определение:

0.3.2. Любая линейно упорядоченная последовательность слов из словаря  $V$  называется цепочкой над словарем  $V$ . Все слова цепочки нумеруются слева направо. Пара объектов слово и номер слова в цепочке называется вхождением слова в цепочку, или, короче, вхождением. Различение слова и вхождения единственный оставленный нами случай различия (ср. § 1) между знаком-типом и знаком-событием.

Понятие «вхождения» при другой терминологии хорошо разъяснено А. И. Смирницким, по которому то, что мы называем вхождениями, является, «нумерически разными случаями употребления этого слова, поскольку каждый такой случай представляет собой эмпирически особый, отдельный акт, отдельное явление, осуществленное . . . в разное время, хотя бы и в непосредственно следующие друг за другом отрезки времени, как, например, в арии Ленского: *Куда, куда, куда вы удалились*<sup>12</sup>. Приведем еще один пример, также использовавшийся А. И. Смирницким.

В цепочке *Глухой<sub>1</sub> глухого<sub>2</sub> свал<sub>3</sub> на<sub>4</sub> суд<sub>5</sub> судь<sub>6</sub> глухого<sub>7</sub>* два разных вхождения слова *глухого*, а именно (*глухого*, 2) и (*глухого*, 7).

## § 17. РЕАЛЬНЫЕ ФРАЗЫ И ЗАПРЕЩЕННЫЕ ЦЕПОЧКИ

В теории лингвистических моделей нас в первую очередь интересуют не произвольно построенные цепочки, а также цепочки, которые соответствуют правильно построенным предложениям некоторого языка.

Примечание. Ранее автор вслед за О. С. Кулагиной<sup>13</sup> употреблял термин «отмеченная фраза». Такое словоупотребление оправдывалось тем, что соответствующие фразы как бы считались отмеченными кем-то со стороны. Отвергнув эту концепцию, мы вынуждены отказаться от ее терминологии. Мы будем употреблять теперь просто термин «фраза» как синоним термина «правильно построенная цепочка».

Важная проблема лингвистической теории моделей состоит в фиксации множества правильно построенных цепочек, т. е. фраз.

До сих пор в теории моделей этот вопрос практически обходили, считая множество правильно построенных, или отмеченных фраз заданными извне. Психологическим оправданием для такого задания, по-видимому, было распространенное мнение,

<sup>12</sup> А. И. Смирницкий. К вопросу о слове. «Труды Ин-та языкознания», IV, 1954, стр. 6.

<sup>13</sup> О. С. Кулагина. Об одном способе определения грамматических понятий на базе теории множеств. «Проблемы кибернетики». М., 1958.

что: а) любой говорящий на данном языке всегда может отличить правильно построенную цепочку от неправильно построенной и б) два разных носителя одного и того же языка дадут одинаковый ответ на вопрос, является ли некоторая предъявленная им цепочка правильно построенной фразой.

В последнее время выяснилось, что обе эти презумпции и, в особенности, вторая из них основаны на довольно шатких основаниях. Так А. Хилл<sup>14</sup> показал, что некоторые фразы, приводимые Хомским<sup>15</sup> как заведомо неправильно построенные, воспринимаются информантом как «поэтические», иногда даже «не только как современная поэзия, но как хорошая современная поэзия» (речь идет, в частности, о цепочке с неправильным порядком слов *Furiously sleep ideas green colorless*).

Имеет место и обратное: по поводу целого ряда фраз из современной поэзии<sup>16</sup> мы с большим трудом можем определить, являются ли они правильно построенными или нет, ср.

Всей бессоницей я тебя люблю (Цветаева)  
... когда поездов расписание  
Камышинской веткой читаешь в купе (Пастернак)  
Пока...  
Я жизнь, как Лермонтова дрожь,  
Как губы в вермут окунал (его же)  
Я учусь словесо (Хлебников)<sup>17</sup>  
Они кажутся засохшее дерево (его же)  
Голод чем меч долог (его же)  
Бобоэби пелись губы (его же)

Некоторые из приведенных выше примеров показывают, что в поэзии появляются цепочки, которые значительным числом говорящих могут восприниматься как неправильно построенные. Во всяком случае далеко не очевидно, какие из приведенных выше поэтических строк будут признаны информантами еще правильно построенными, а какие уже неправильно построенными.

Поэтому, если включать в рассмотрение и поэтические факты, то вряд ли правильно было бы предполагать, что любой носитель

<sup>14</sup> См.: А. Хилл. О грамматической отмеченности предложения. Перев. с англ. — ВЯ, 1962, № 4.

<sup>15</sup> См.: Н. Хомский. Синтаксические структуры. «Новое в лингвистике», вып. II. М., 1962.

<sup>16</sup> О значении поэтических фактов для решения вопроса о грамматической правильности, см.: S. R. Levin. Poetry and grammaticality. «Proceedings of the ninth International congress of linguists». The Hague, 1964, стр. 308—315; см. также: Ю. И. Левин. О некоторых чертах плана содержания в поэтических текстах. «Структурная типология языков». М., 1966, стр. 212.

<sup>17</sup> Ср. анализ грамматической структуры хлебниковских строк в работе: Р. Якобсон. Новейшая русская поэзия, набросок первый. Прага, 1921.

языка располагает алгоритмом, распознающим грамматическую правильность произвольной цепочки.

Кроме того, существенно отметить, что многие фразы типа *идея яростно спит* воспринимаются как семантически неправильные (см. § 61) по сравнению с фразами типа *девочка спокойно спит*, как в каком-то смысле «отклоняющиеся» (deviant)<sup>18</sup>, причем Путнам правильно подчеркивает<sup>19</sup>, что одна из задач лингвистической теории состоит в том, чтобы как раз объяснить, в чем состоит «отклонение» и как понимаются носителем языка «отклоняющиеся» цепочки.

Между тем ясно, что априорное задание всех правильно построенных, или потенциальных, фраз, игнорирует эту важнейшую проблему.

Однако наиболее трудным логическим моментом в априорном задании множества фраз является то, что такое исходное понятие любой лингвистической модели задается на основании трудно оправдываемых предпосылок об интуиции человека, говорящего на данном языке, ибо приведенные выше факты свидетельствуют о том, что эта интуиция далеко неоднозначна. Можно было бы, правда, считать, что имеется всего один «оракул», к которому разные исследователи всегда могут обратиться и который в случае сомнения дает однозначный ответ на вопрос, является ли цепочка фразой или нет<sup>20</sup>. К сожалению, этот прием является фикцией.

Дело в том, что одна из целей метода лингвистических моделей состоит в том, чтобы предложить такую теорию, пользуясь которой два разных исследователя, располагающие той же самой совокупностью исходных объектов, должны были бы прийти к одинаковым выводам. Для этого надо иметь возможность точно зафиксировать совокупность исходных объектов и в первую очередь множество правильно построенных фраз. Но особенность естественных языков состоит в том, что это множество практически (а в любой разумной модели и теоретически) бесконечно и поэтому не может быть задано списком (или в случае оракула не может быть получено от него в какой-то ограниченный промежуток времени).

Итак дело не только в обращении к интуиции, но и в том, что все фразы даже не могут быть заданы как объективная данность. Другое дело совокупность реально встретившихся исследователю (скажем, в данный отрезок времени) фраз, например совокупность всех фраз данной книги или всех книг данного писа-

<sup>18</sup> Х. Путнам. Некоторые спорные вопросы теории грамматики. Перев. с англ. «Новое в лингвистике», вып. IV. М., 1965.

<sup>19</sup> Путнам ссылается при этом на работу Зиффа «Семантический анализ» (P. Ziff. Semantic analysis. Ithaca, 1960).

<sup>20</sup> Именно такова была аргументация в кн.: «Модели языка» (стр. 89).

теля или данного списка писателей. При таком задании исходного объекта можно быть уверенным, что два исследователя будут пользоваться одинаковым исходным объектом или, иначе говоря, что объект определен однозначно.

В § 19 и § 20 мы предложим процедуру, исходя из объективной данности, а именно из множества фраз, реально встретившихся исследователю, выводит все те цепочки, которые мы хотели бы считать правильно построенными.

Эти процедуры связаны с порождением — на основе конечной фиксированной совокупности объектов — бесконечного множества фраз и устроены так, что должны достаточно простым способом породить все заданные реальные фразы, а также еще какие-то, которые и считаются правильно построенными.

При таком подходе, однако, необходимо иметь заданным и некоторое множество заведомо неправильных, или «запрещенных», цепочек. В самом деле, если такое множество запрещенных цепочек не задано, то простейшей порождающей процедурой, дающей все реальные фразы, будет следующее простое предписание: «состав произвольную последовательность букв (или слов) данного алфавита (или словаря)». Как знали еще лапутянские академики у Свифта, устройство, работающее по данному предписанию, может породить, между прочим, и все реальные фразы, которые когда-либо были (и даже могли бы быть) написаны. Но неправильные цепочки, порождаемые таким устройством (а их подавляющее большинство), лишают модель всякой лингвистической ценности.

Для того, чтобы порождаемые нереальные цепочки могли соответствовать некоторому лингвистически значимому понятию, нужно, чтобы порождающее устройство порождало некоторые заведомо правильные («реальные») фразы и не порождало ни одной заведомо неправильной цепочки.

Отсюда необходимость ввести еще понятие «множество запрещенных цепочек». Практически это означает, что моделирование начинается на более высоком уровне знания языка, чем тот, который предполагается, например, в дескриптивной лингвистике. Мы предполагаем, что исследователю знакомы уже некоторые правила языка, нарушение которых недопустимо. Такие правила представляют собой, однако, такую же объективную данность, как и реальные фразы (такие правила могут быть даны конечным обозримым списком). Больше того, интуиция всех говорящих дает однозначный ответ на вопросы типа, можно ли сказать по-русски *я учимся* или *ты учусь*.

Здесь нужно иметь в виду, что мы вовсе не требуем дихотомии «правильно построенные цепочки» — «неправильно построенные цепочки». Важно лишь то, что нам задано некоторое множество реальных фраз и некоторое множество запрещенных цепочек и что эти множества не пересекаются.

Заметим, что именно таким был первоначальный подход Хомского, который писал: «Для содержательной постановки задач грамматики достаточно предположить лишь частичное знание предложений и непредложений. Это значит, что в рамках данного рассмотрения мы можем допустить, что некоторые последовательности фонем являются определенно непредложениями. Во многих промежуточных случаях мы должны предоставить самой грамматике решать вопрос о грамматической правильности предложения, если грамматика построена простейшим образом так, что в нее включаются несомненные предложения и исключаются несомненные непредложения»<sup>21</sup>.

К сожалению, в дальнейшем Хомский дал повод к мнению, что для его концепции существенна жесткая дихотомия «грамматически правильных» и «грамматически неправильных» цепочек. Поэтому критические замечания Хилла оказались вполне справедливыми<sup>22</sup>.

В западноевропейской литературе по теории моделей выдвинута одна интересная концепция, близкая в некоторой степени к предлагаемой нами. Речь идет о модели Паркер-Родса<sup>23</sup>, в которой дихотомия «высказывания» (utterances) — «не-высказывания», т. е. «правильно построенные» — «неправильно построенные цепочки» принимается лишь для «закрытого языка», т. е. языка, содержащего конечное множество фраз. Что же касается «открытого языка», то для него рассматривается четыре разряда цепочек.

1. Цепочки, которые любым носителем языка воспринимаются как правильно построенные фразы, например, *Я гуляю по парку*.

2. Цепочки, которые любым носителем языка воспринимаются как неправильно построенные, но которые легко могут быть этим носителем языка исправлены, например, цепочка *\*Пароход плыла по река*, по поводу которой любой носитель языка может заметить: надо сказать *пароход плыл по реке*.

3. Цепочки, которые воспринимаются как неправильно построенные, но по поводу которых нельзя с определенностью сказать, как их надо исправить: *\*три едят медведя, но собаку*.

4. Цепочки, по поводу которых большинство носителей языка не может утверждать, ни что они правильны, ни что они неправильны, поскольку они не знают всех входящих в их состав

<sup>21</sup> Н. Хомский. Синтаксические структуры, стр. 147.

<sup>22</sup> Ср. также: А. V. Isačenko. Gramatičnost a význam. «Acta Universitatis Carolinae. Philologia, Slavica Pragensia», IV, 1962, стр. 48.

<sup>23</sup> А. F. Parker-Rhodes. A new model of syntactic description. «1961. International Conference on Machine Translation of languages and Applied Language Analysis», vol. 1. London, 1962, стр. 29—32.

слов, например, макаронические стихи типа Мятлевских:

иль достанет а креди  
это русских малади...  
...тут одетые сорочкой,  
или блузой, два коми,  
первые де мез-ами

Паркер-Родс вводит одно определение, которое мы приспособим здесь к нашим терминам: Цепочка  $f$  в открытом языке  $L$  называется и с п р а в и м о й, если она не есть реальная фраза и отличается от некоторой реальной фразы  $f'$  таким образом, что любой компетентный носитель языка однозначно идентифицирует  $f$  с  $f'$ .

Но множество «исправимых цепочек» включает все цепочки, которые мы хотели бы интерпретировать в качестве «запрещенных». Таким образом, мы видим, что в разных концепциях постепенно осознается необходимость ввести в рассмотрение некоторое множество цепочек заведомо неправильных (именно потому, что ясно, как их исправить). В советской литературе по лингвистическим моделям близкую к концепции автора точку зрения (хотя в деталях и отличную от нее) высказал М. И. Белецкий, который писал: «Предположение о том, что множество последовательностей естественно делится на грамматически правильные и грамматически неправильные, оказалось неверным. Мы можем сказать только, что некоторые последовательности (*идея яростно спит*) являются заведомо грамматически правильными, некоторые (*лодка в шлюпке лодка была течь*) заведомо неправильными; кроме того, имеется много последовательностей, о которых мы не можем с полной определенностью сказать ни того, ни другого. Относительно последних мы иногда только можем сказать, что одни из них представляются нам более (или менее) правильными, чем другие»<sup>24</sup>.

Далее Белецкий предлагает описывать меру правильности в вероятностных терминах. Сходные идеи были высказаны автором в другом месте<sup>25</sup>. В данной работе мы, однако, не включаем в рассмотрение вероятностный аспект, поэтому в дальнейшем мы предложим лишь логический критерий для степени правильности (см. § 19).

Близка к концепции автора с идейной точки зрения, хотя и сильно отличается от нее по практическим выводам, концепция Б. А. Успенского<sup>26</sup>, который также исходит из задания реальных

<sup>24</sup> М. И. Белецкий. Модель русского языка. «Научно-техническая информация», 1964, № 7, стр. 41.

<sup>25</sup> В рецензии на сб. «*Studia grammatica*» (см.: ВЯ, 1964, № 3).

<sup>26</sup> Б. А. Успенский. Структурная типология языков. М., 1965, стр. 75—76.

фраз, которые он называет реально правильными предложениями («имеется в виду правильность во всех отношениях: и семантическая и грамматическая и др.»<sup>27</sup>). Б. А. Успенский определяет на основе этого понятия новое понятие грамматически правильного предложения, а именно грамматически правильным является всякое предложение, «которое можно превратить в реально правильное путем замены каких-то корневых элементов в этом предложении»<sup>28</sup>. Как указывает автор этой концепции, она приводит к одному довольно неожиданному выводу, а именно в таких языках, как все славянские (в частности, русский язык), следует считать, что каждое слово имеет все три рода, т. е. считать, например, что предложение \**вижу маленькую стулу* грамматически правильно, ибо оно получено заменой корневого элемента *женщин-* на корневой элемент *стул* в реально правильном предложении *вижу маленькую женщину*. Такой подход на самом деле вполне последователен и имеет определенный смысл для тех целей, которые поставил себе данный автор, но для наших целей этот подход неприемлем, ибо, во-первых, целый ряд понятий, вводимых нами в дальнейшем, потерял бы всякий смысл (правда, это может быть истолковано как показатель бессмысленности этих понятий) и, во-вторых, — что уже важнее, — мы не могли бы предложить определений для таких категорий, как категория грамматического рода (и ряд других категорий, обладающих общими формальными свойствами с этой категорией, — см. главу VI), а, с другой стороны, не смогли бы определить в своих терминах отличие согласования от управления<sup>29</sup>.

По-видимому, уровень грамматической правильности, рассматриваемый Б. А. Успенским, полезен для общей типологии языков мира, ибо на этом уровне хорошо выявляются различия между далекими по структуре языками. Однако для близкородственных языков, какими являются славянские, мы потеряли бы, перейдя на этот уровень, ряд существенных различий между этими языками, связанных с типологией рода.

Показательно, однако, что в концепции Б. А. Успенского вместо недифференцированного и неконструктивного понятия правильности появилось два дифференцированных и точно определенных понятия: реальной правильности и грамматической правильности. Поэтому мы и говорим об идейной близости концепции Б. А. Успенского и концепции, предлагаемой нами далее в §§ 19—20 (заметим, что идея Б. А. Успенского частично отражена нами в § 21 в понятии «обратно-алфавитного разбиения», которое для славянских языков дает близкую картину).

<sup>27</sup> Б. А. Успенский. Структурная типология языков. М., 1965, стр. 75.

<sup>28</sup> Там же, стр. 76.

<sup>29</sup> Характерно, что в концепции Б. А. Успенского нет существенного различия между согласованием и управлением (см.: Указ. соч., стр. 128).



Наконец, в последних работах чешского математика Л. Небеского<sup>30</sup> также выдвигается идея порождения множества грамматически правильных фраз из некоторой совокупности данных фраз, которые он называет осмысленными (*smysluplné*). Небеский заметил, что если осмыслены три фразы типа *мальчик читает*, *мальчик стоит*, *стол стоит*, то фраза *стол читает* будет грамматически правильной. В связи с этим он предлагает говорить, что если для любых  $f_1, f_2$  и  $f_3$  и любых слов  $x, x', y, y'$  фразы  $f_1x_2yf_3$ ,  $f_1x'f_2yf_3$  и  $f_1xf_2y'f_3$  осмысленны, то цепочка  $f_1x'f_2y'f_3$  грамматически правильна в первом приближении. Оговорка о первом приближении сделана потому, что таким способом можно породить и грамматически неправильные цепочки, например, если *читает* заменить на *читал*, а *стол* на *лампа*, ибо фразы *мальчик стоит*, *лампа стоит*, *мальчик стоял* осмыслены, а цепочка *лампа стоял* — грамматически неправильна. Но Небеский замечает, что если взяты не четверки, а шестерки слов, рассматривая фразы вида  $f_1x_2yf_3z_4f_4$  и т. п., то соответствующее определение уже вполне хорошо описывает ситуацию в реальных языках.

Мы не будем заниматься здесь обобщением модели, рассматриваемой Небеским, и вводимым в связи с этим понятием типа языка, ибо нам важно подчеркнуть лишь основную идею порождения одних грамматически правильных фраз на базе других.

Итак, вместо задания множества «отмеченных фраз» мы перешли к другой концепции, как представляется, более гибкой и более соответствующей лингвистической практике: заданы два конечных и непересекающихся множества реальных и запрещенных цепочек. Это удобно и по следующим соображениям. В зависимости от выбранного множества реальных фраз будет меняться, разумеется, и все множество фраз. Такое изменение множества реальных фраз может определяться целью, поставленной себе исследователем (а также степенью овладения исследуемым языком). Эта процедура не снижает строгости теории, если точно указывается множество реальных, а в дальнейшем и правильно построенных фраз, на которых выполняются операции, предусмотренные моделью.

Таким образом, создается возможность подходить к понятию фразы как к понятию экспериментальному.

Предложив менять в зависимости от цели исследования множество реальных фраз, каждый раз фиксируя это множество, мы в сущности перешли к гипотетико-дедуктивной точке зрения в теории моделей, на которой настаивал С. К. Шаумян<sup>31</sup>.

<sup>30</sup> L. Nebeský. K pojmu věty smysluplné a věty gramaticky správné. «Kybernetika a její využívaní». Praha, 1965, стр. 296—298. Первый набросок этой идеи содержится уже в работе: Он же. O jednom algebraickem modelu jazyka. «Slovo a slovesnost», 1963, № 4, стр. 231—237.

<sup>31</sup> С. К. Шаумян. Проблемы теоретической фонологии. М., 1962.

Исходные понятия подбираются путем ряда экспериментов, так, чтобы определить в конце концов искомые лингвистические явления. Этот подход столь же научен, как и чисто дедуктивный подход, и гораздо более подходит к описанию сложного материала, логическая структура которого целиком в интуиции говорящих. Суть метода состоит в том, чтобы для каждого важного лингвистического понятия, например, «часть речи» или «падеж» или «род» найти такую совокупность исходных понятий и такую операционную процедуру, чтобы в результате данное понятие выводилось достаточно строго.

## § 18. СИНТАКСИЧЕСКАЯ ЗАВИСИМОСТЬ И СМЫСЛ ФРАЗЫ

В ряде исследований, посвященных синтагматическим моделям, для каждой фразы считается заданным отношение непосредственного синтаксического подчинения  $D$ , определенное на множестве вхождений в эту фразу<sup>32</sup>. Иначе говоря, для каждого слова, входящего во фразу, считается известным слово, от которого данное слово непосредственно зависит.

Мы будем пользоваться следующим способом графического представления отношения  $D$ : если вхождение  $x_j$  непосредственно подчинено вхождению  $x_i$ , то мы на чертеже проведем стрелку от  $j$  к  $i$ .

Обычно расстановка стрелок непосредственного подчинения целиком соответствует интуиции, например, во фразе  $y_1$  *вижу*<sub>2</sub> *маленькую*<sub>3</sub> *девочку*<sub>4</sub> стрелки расставляются следующим способом.



Здесь единственным пунктом, могущим вызвать сомнение, является соглашение, что в группе  $y$  *вижу* слово  $y$  считается непосредственно подчиненным слову *вижу*. Причина этого будет указана несколько ниже.

Пожалуй, при задании зависимостей между полнозначными словами — это единственный случай, вызывающий споры. Зато

<sup>32</sup> Формальный анализ этого отношения см.: С. Я. Фитиалов. О моделировании синтаксиса в структурной лингвистике. «Проблемы структурной лингвистики». М., 1962, стр. 110—114; М. И. Белецкий, В. М. Григорян, И. Д. Заславский. Аксиоматическое описание порядка и управления слов в некоторых типах предложений. «Математические вопросы кибернетики и вычислительной техники». Ереван, 1963, стр. 71—85; P. Novák. Nekteré otázky syntaktické analyzy (z hlediska strojového překladu). «Slovo a slovesnost», 1962, № 1, стр. 11—12.

когда устанавливаются зависимости между служебными и полнозначными словами, то возникают довольно произвольные решения<sup>33</sup>.

Мы будем поэтому учитывать, что отношение  $D$  распадается в каждой фразе на два отношения:

$D^{(полн)}$  — отношения между полнозначными словами,

$D^{(служ)}$  — отношения между служебными и полнозначными словами.

Различие между этими двумя типами отношений отражено и в традиционной теории словосочетаний (словосочетаниями как правило считаются лишь пары связанных между собой полнозначных слов).

Что касается отношения  $D^{(служ)}$ , то мы будем считать, что оно устанавливается следующим образом:

а) предлог непосредственно подчиняет зависящее от него существительное, например, *перед* непосредственно подчиняет *домом*;

б) личная форма глагола непосредственно подчиняет неличную, например, *буду* непосредственно подчиняет *читать*;

в) союз однородности, например, *и* непосредственно подчиняет связываемые им слова, например, в группе *вижу Ваню и Петю* слово *вижу* непосредственно подчиняет слово *и*, а слово *и* непосредственно подчиняет слово *Ваню* и слово *Петю*. Этот же союз непосредственно подчиняет слова, определяющие все однородные члены, связываемые им, например, в группе *маленькие мальчики и девочки* мы принимаем, что союз *и* непосредственно подчиняет слово *маленькие*. Идея такой расстановки стрелок заимствована нами из работы Ю. А. Шрейдера<sup>34</sup>. Она сводится к тому, что группа однородных членов трактуется как одно слово, а представителем этого «слова» выступает союз однородности. Полный смысл этого соглашения станет ясным в § 56, но некоторые основания для такой расстановки стрелок синтаксического подчинения можно привести уже сейчас.

Обычно принимается (и мы будем в книге, по возможности, следовать этому допущению), что граф отношения непосредственного подчинения должен представлять собой дерево. Именно в связи с этим оказывается удобным считать, что личная форма глагола является корнем дерева (в случае нескольких однородных личных форм корнем дерева по принятому соглашению будет «представляющий» их союз).

---

<sup>33</sup> Довольно распространенный способ представления синтаксических зависимостей см.: И. А. Мельчук. Автоматический синтаксический анализ, т. I. Новосибирск, 1964, стр. 20—25.

<sup>34</sup> Ю. А. Шрейдер. Свойства проективности языка. «Научно-техническая информация», № 8, 1964, стр. 41 (Ю. А. Шрейдер в свою очередь ссылается на неопубликованный доклад Е. В. Падучевой).

Но при более естественной расстановке стрелок непосредственного подчинения, например, если считать, что слово *маленькие* в приведенном примере непосредственно подчинено слову *мальчики* и непосредственно подчинено слову *девочки*, нарушаются аксиомы дерева<sup>35</sup>.

Приведенная система расстановки стрелок с интуитивной точки зрения также достаточно неудовлетворительна, но этим будет отличаться в данном случае л ю б а я расстановка стрелок, ибо здесь мы имеем дело с с о ч и н е н и е м и поэтому л ю б о й способ представления связи как п о д ч и н и т е л ь н о й будет неадекватным. В главе VII мы увидим, однако, что предложенный здесь способ, при котором сочинительная группа предстанет как один целый элемент, удобен для решения ряда вопросов (см. § 56).

Наряду с отношением *D* нам понадобится еще одно отношение:

0. 3. 3. Отношением синтаксического подчинения называется транзитивное замыкание отношения непосредственного подчинения.

Каково же место отношения *D* (и соответственно *D*<sup>+</sup>) в общей системе понятий, которые отражают структуру языка?

Постепенно в теории моделей становится все более ясным, что в числе исходных понятий теории, если она хочет охватить достаточно широкий класс фактов, обязательно должно присутствовать хотя бы одно из трех следующих понятий:

а) вместе с множеством фраз должна быть задана грамматика, порождающая эти фразы (о понятии грамматики см. § 20);

б) вместе с множеством фраз должна быть задана для каждой фразы система синтаксических отношений;

в) вместе с множеством фраз должно быть задано для каждой фразы определенное отношение, отражающее смысл фразы.

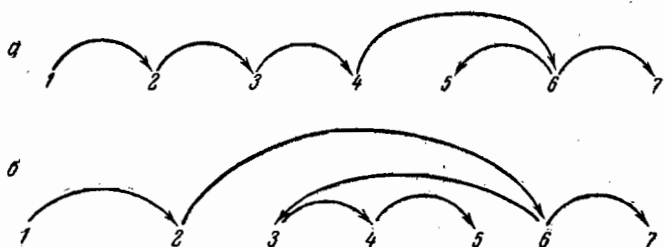
Мы вернемся к связи соответствующих концепций в главе VII, а здесь изберем последнюю точку зрения.

Прежде всего заметим, что задание смысла не исчерпывается заданием парадигм, т. е., в частности, соответствия между словами и объектами действительности, ибо одна и та же фраза, как известно, может передавать принципиально разные содержания. Это имеет место при всех случаях так называемой конструктивной омонимии, например:

*при*<sub>1</sub> *условии*<sub>2</sub> *соблюдения*<sub>3</sub> *принципов*<sub>4</sub> *взаимности*<sub>5</sub> *и*<sub>6</sub> *выполнения*<sub>7</sub> (*взятых на себя обязательств*), где речь может идти а) о принципе взаимности и принципе выполнения взятых на себя обязательств и б) об условии соблюдения одного и выполнения другого.

<sup>35</sup> В ряде работ (например: И. А. М е л ь ч у к. Указ. соч.) приняты иные соглашения, также обеспечивающие в подобных случаях выполнение аксиом дерева, но выход, предложенный Падучевой и Шрейдером, нам кажется более удобным.

На лингвистическом уровне мы здесь имеем дело с разн<sup>ой</sup> системой синтаксических зависимостей:



Ясно, что если для каждого предложения была бы задана система синтаксических зависимостей, то все подобные проблемы было бы очень легко разрешить.

Как мы покажем в § 55, для реальных фраз систему синтаксических зависимостей можно получить из более примитивного понятия, а именно из задания на множестве реальных фраз отношения: «смысл фразы *f* включает в себя смысл фразы *g*». Заметим, что отсюда можно уже вывести отношение тождества смысла.

На первый взгляд включение понятия смысла делает теорию неформальной и неточной. Такого рода опасения заставляли многих исследователей отказаться от понятий, связанных со смыслом (во всяком случае для автора такого рода опасения долгое время были решающими). Конечно, система понятий, использующая только формальные признаки лингвистических объектов, обладает высокой степенью привлекательности. Однако без задания разбиения на парадигмы, мы можем построить лишь довольно бедную теорию, не отражающую ряд значительных интуитивных понятий (в частности, столь важного для славянских языков понятия «части речи»).

Введя понятие парадигмы, мы в теории моделей формальным путем вводим понятие одинаковости лексического значения и получаем лингвистически важные результаты.

Почему же нельзя подобным путем ввести понятие одинаковости смысла (тем более, раз выяснилось, что смысл фразы не исчерпывается перечислением объектов и процессов действительности, о которых идет речь)?

До сих пор этому мешало одно существенное различие между заданием одинаковости лексического значения слов и заданием одинаковости смысла фраз. Число слов как правило предполагается конечным. Поэтому задание парадигм может быть практически осуществлено списком пар всех словоформ. Что же касается лингвистически значимых фраз, то до сих пор это понятие моделировалось, как мы знаем, при помощи понятия грамматиче-

ской правильности, причем в большинстве моделей множество грамматически правильных фраз считалось бесконечным. При таком задании, любой вопрос о задании отношения одинаковости смысла на множестве всех фраз становился бессодержательным.

Но раз мы задаем конечное (притом практически обозримое) множество реальных фраз, то на этом множестве отношение одинаковости смысла вполне может быть задано формально как список пар всех реальных фраз с указанием, находятся ли они в данном отношении или нет.

Итак, мы считаем заданным:

а) множество словоформ, т. е. словарь  $V = \{x\}$  и разбиение его на парадигмы;

б) систему категорий и принадлежность каждого слова к определенному категориальному пучку;

в) множество реальных фраз  $\alpha$ , причем каждое слово входит хотя бы в одну реальную фразу;

г) множество запрещенных цепочек  $\beta$ ;

д) отношение: смысл фразы  $f$  включает в себя смысл фразы  $g$ , определяемое на множестве реальных фраз.

#### § 19. ОДИН СПОСОБ ПОРОЖДЕНИЯ ФРАЗ НА ОСНОВЕ РЕАЛЬНЫХ И ЗАПРЕЩЕННЫХ ЦЕПОЧЕК

0.3.4. Будем говорить, что слово  $x$  находится в отношении дистрибутивной близости  $P$  с  $y$  (записывается  $xPy$ ), если выполнены следующие два условия:

а) существуют цепочки  $f$  и  $g$ , такие, что  $fxg$  и  $fyg$  суть реальные фразы и

б) не существует  $f$  и  $g$ , таких, что из цепочек  $fxg$  и  $fyg$  одна относится к реальным фразам, а другая — к запрещенным цепочкам.

Отношение дистрибутивной близости  $P$  рефлексивно и симметрично, но не транзитивно. Поэтому мы будем пользоваться в некоторых случаях (в особенности в гл. VII) наряду с отношением  $P$ , также и транзитивным замыканием этого отношения, и обозначим это новое отношение через  $Q$ .

Таким образом, имеет место  $xQy$ , если существует  $n$  слов  $z_1, \dots, z_n$ , таких, что  $x = z_1$ ,  $y = z_n$  и для любого  $i$  верно  $z_i P z_{i-1}$ .

0.3.5. Сегментом, развертывающим слово  $x$ , называется цепочка  $Y$ , состоящая ровно из двух слов и обладающая тем свойством, что

а) существуют цепочки  $f$  и  $g$ , такие, что  $fxg$  и  $fYg$  суть реальные фразы и

б) не найдется  $f$  и  $g$ , таких, что цепочка  $fxg$  есть реальная фраза, а  $fYg$  относится к запрещенным цепочкам.

0. 3. 6. Назовем сегмент *регулярным*, если среди слов, которые он развертывает, находится правое или левое слово сегмента. В этом случае развертываемое слово называется *ядром* сегмента, а оставшаяся часть *определителем*.

Исходя из отношения *P* и понятия сегмента, мы дадим теперь следующее определение множества фраз  $\theta$ .

0. 3. 7. Множество фраз  $\theta$  формируется следующим образом:

- а) всякая реальная фраза есть фраза;
- б) цепочка есть фраза, если она получена из фразы путем замены некоторого слова *x* на слово *y*, такое, что *xPy*;
- в) цепочка есть фраза, если она получена из фразы путем замены некоторого слова *x* развертывающим его сегментом *Y*;
- г) других фраз нет.

При порождении фраз слово может заменяться сегментом, но обратное не допускается. Это ограничение введено сознательно. Связано это с тем, что произвольная замена сегмента на слово и слова на сегмент могла бы привести к тому, что в процессе получения фраз мы переходили бы не только от более коротких к более длинным, но и обратно от более длинных к более коротким. В этом случае множество фраз могло бы оказаться неразрешимым (см. § 11), т. е. нельзя было бы вообще говоря, построить алгоритм, распознающий, произвольная цепочка есть фраза или нет. Наш способ порождения, при котором не получается более коротких фраз, гарантирует разрешимость множества фраз<sup>36</sup>.

Перейдем теперь к одному конкретному примеру. Пусть реальные фразы суть:

- 1) *девочка спокойно спит*
- 2) *мальчик равнодушно говорит*
- 3) *мальчик спокойно говорит*
- 4) *равнодушно взглянуть*
- 5) *яростно взглянуть*
- 6) *счастливая идея*
- 7) *счастливая девочка*
- 8) *девочка спит*
- 9) *красивая девочка спокойно спит*
- 10) *очень красивая девочка спокойно спит*
- 11) *мальчик спит*
- 12) *счастливая девочка спит*
- 13) *идея красоты*

Пусть запрещены цепочки

- 14) *\*счастливая девочка спит спит*
- 15) *\*счастливая мальчик*
- 16) *\*девочка спит красоты*

<sup>36</sup> Это следует хотя бы из рассуждений в цитированной статье Путнама.

Тогда *равнодушно* дистрибутивно близко к *спокойно* (ср. фразы 2 и 3); *равнодушно* дистрибутивно близко к *яростно* (ср. фразы 4 и 5); *идея* дистрибутивно близко с *девочка* (ср. 6 и 7); *очень красивая* — регулярный сегмент, развертывающий слово *красивая* (ср. фразы 9 и 10); *красивая девочка* — регулярный сегмент, развертывающий слово; *девочка* (ср. фразы 1 и 9), *спокойно спит* — регулярный сегмент, развертывающий слово *спит* (ср. фразы 1 и 8).

Примечание. От процедуры, предложенной в «Моделях языка» (стр. 65), данная процедура отличается тем, что ей практически обеспечивается (благодаря введению понятия запрещенной цепочки) непорождение цепочек вида *счастливая мальчик*. В самом деле, хотя *мальчик* и *девочка* встречаются в одинаковом окружении (фразы 8 и 11), имеются цепочки 7 и 15, одна из которых реальная фраза, а вторая запрещена. Заметим также, что *девочка спит* не есть сегмент, развертывающий слово *девочка*, ибо хотя реальные фразы 7 и 12, имеется реальная фраза 12 и запрещенная цепочка 14, или развертывающий слово *идея*, ибо хотя 6 и 12 реальные фразы, имеется реальная фраза 13 и запрещенная цепочка 16.

Эта процедура позволяет получать из конечного множества заданных фраз бесконечное множество, причем это обеспечивается регулярностью соответствующих сегментов. Так, например, все цепочки вида: *очень... очень красивая... красивая девочка спокойно... спокойно спит*, — суть фразы.

Покажем теперь, что в нашем примере цепочка *идея яростно спит* есть фраза. Цепочка 1 есть фраза по пункту а) определения 4. 5. По пункту б) суть фразы: *девочка яростно спит*, *идея яростно спит*.

Теоретически возможность порождения данной моделью запрещенных цепочек остается:

Т. 4. 1. а) Процедурой данного параграфа можно породить, исходя из реальных и запрещенных цепочек в числе фраз, запрещенную цепочку и б) существуют цепочки, которые не запрещены и не являются фразами.

Вторая из этих особенностей лингвистически вполне оправдана, она объясняет возможность появления в стихах в качестве поэтически оправданных (ср. § 61) таких цепочек как приведенные на стр. 66 хлебниковские строки, которые никак нельзя объяснить аналогией в отличие, например, от фразы *всей бессоницей я тебя люблю*, понимаемой именно по аналогии с фразами типа *всей душой я тебя люблю*<sup>37</sup>.

Но если так интерпретировать нашу процедуру, то становится очевидным, что модель здесь содержит аналог резервуара, который необходим как для развития повседневного языка, так и для поэтического творчества<sup>38</sup>.

<sup>37</sup> Ср. также объяснение Путнамом процесса понимания фраз типа стиха Дилана Томаса *A grief ago I saw him there*, который можно было бы перевести так: 'Печаль тому назад я видел его там', ср. у Рильке: *Nur ein Lächeln lang* («Stundenbuch»).

<sup>38</sup> Ср.: S. R. Levin. Poetry and grammaticality.



Следует отметить, что процедуры этого параграфа, а также следующего не предусматривают получения фраз путем применения правил трансформации. Тем самым остается возможность обогащать множество фраз не только процедурами, рассматриваемыми здесь, но и всякого рода трансформациями, включающими как достаточно регулярные, так и индивидуальные преобразования (ср. § 51).

Что же касается первой особенности, то она, несомненно, свидетельствует о логическом несовершенстве предложенной процедуры. Можно было бы так усложнить процедуру (требуя, например, проверки при каждой замене), чтобы среди фраз не появилось ни одной запрещенной цепочки. Но дело в том, что для естественных языков предложенная простая процедура работает достаточно хорошо. Возможно, что отличительная черта естественных языков (в отличие от абстрактных языков вроде того, который был использован при доказательстве теоремы 4. 1) в том, что в них фраза не может совпасть с запрещенной цепочкой<sup>39</sup>. С другой стороны, не исключено, что и в естественных языках фразы на некотором этапе развития языка могут пересекаться с запрещенными цепочками, но тогда следует считать, что действует чисто стилистический закон, по которому в таком случае происходит перераспределение запрещенных цепочек или реальных фраз, которое ликвидирует подобное совпадение.

Интерес предложенной процедуры и в том, что позволяет поставить вопрос об «уровнях правильности». Можно выделять разные «уровни правильности», ограничивая число производимых замен<sup>40</sup>. С этой точки зрения фраза *идея равнодушно спит* будет «более правильной», чем *идея яростно спит*, а *девочка равнодушно спит* правильнее, чем *идея равнодушно спит*. Фраза *очень очень очень красивая девочка* окажется менее правильной, чем фраза *очень очень красивая девочка*.

Все это показывает, что подобные процедуры смогут, по-видимому, занять в теории лингвистических моделей место, которое в традиционном языковедении занимает (или должна занимать) стилистика (та часть ее, которая относится к культуре речи). Введение в семиотику специальной науки, изучающей, какие фразы можно считать правильно построенными, было предложено Карри<sup>41</sup>, который выделил специальную отрасль

<sup>39</sup> На это обстоятельство обратил внимание автора В. А. Успенский.

<sup>40</sup> Указано автору Ю. А. Шпейдером.

<sup>41</sup> H. B. C u r r y. Some Logical Aspects of Grammatical Structure. «Structure of language and its mathematical aspects». — PSAM, v. 12. Providence, 1961, стр. 59. — Мы отходим здесь от принятого нами принципа ссылаться в тех случаях, когда не важен год появления работы, только на существующий русский перевод (не упоминая оригинал). Дело в том, что в русском переводе («Новое в лингвистике», вып. IV, стр. 102) соответствующее место почти целиком обесмысленно употреблением термина «грамматика» (вместо «учение о грамматичности» или чего-нибудь в этом роде).

семантики, названную им *grammatics* (в отличие от *grammar*).

Аналогии в постановке задач такой науки, с одной стороны, и стилистики, с другой стороны, достаточно глубоки. В самом деле, стилистика в сущности учит (или должна учить), как строить фразы, подобные тем, которые есть в произведениях классической литературы («реальные фразы» нашей модели) и, с другой стороны, трактаты по культуре речи представляют собой собрания фраз, недопустимых по тем или иным основаниям («запрещенные цепочки»).

Можно думать, что осмысление некоторых проблем и подходов стилистики будет полезным для дальнейшего развития теории моделей.

## § 20. ПОНЯТИЯ ГРАММАТИКИ И МНОЖЕСТВА ФРАЗ, ПОРОЖДАЕМЫХ ГРАММАТИКОЙ

Термин «грамматика» в лингвистике употребляется достаточно расплывчато. В теории лингвистических моделей предложено следующее уточнение одного из употреблений этого термина.

0.3.8. Пусть задано: а) множество  $V$ , именуемое «основным словарем», б) множество  $V_1$ , именуемое «вспомогательным словарем», в) котором выделяется специальный элемент «Предложение», в) совокупность  $S$  правил вида  $A \rightarrow B$ , где  $A$  и  $B$  суть цепочки над  $V \cup V_1$ , а символ  $\rightarrow$  не принадлежит  $V \cup V_1$ . Тогда совокупность четырех объектов  $G = (V, V_1, \text{Предложение}, S)$  называется грамматикой.

Правила  $A \rightarrow B$  интерпретируются как разрешения заменить цепочку  $A$  цепочкой  $B$ .

Основной словарь  $V$  имеет интерпретацию, обсуждавшуюся в § 14 (т. е. элементами его являются словоформы, возможна, однако, и такая интерпретация, где элементами  $V$  выступают морфемы, именно такова интерпретация, представленная у Хомского)<sup>42</sup>. В теории грамматик элементы  $V$  принято называть терминальными символами.

Вспомогательный словарь состоит из имен грамматических категорий, поставленных в соответствие не только отдельным словам (такие категории мы обсуждали в § 15), но и группам слов (так называемые синтаксические категории)<sup>43</sup>.

Механизм действия грамматики описывается следующими определениями:

<sup>42</sup> Н. Хомский. Синтаксические структуры.

<sup>43</sup> До сих пор мы не могли обсуждать таких категорий, как сопоставленных отдельным неединичным цепочкам, ибо число таких цепочек бесконечно и принадлежность каждой цепочки к определенной синтаксической категории не могла быть задана разумным способом.

**0. 3. 9.** Говорят, что цепочка  $w$  непосредственно выводима из цепочки  $u$ , если  $u$  имеет вид  $\chi_1 A \chi_2$ ,  $w$  имеет вид  $\chi_1 B \chi_2$ , где  $\chi_1$  и  $\chi_2$  произвольные цепочки над  $V \cup V_1$ , а правило  $A \rightarrow B$  входит в  $S$ . Говорят, что  $w$  выводима из  $u$ , если существует  $n$  цепочек  $v_1, \dots, v_n$  таких, что  $v_1 = u$ ,  $v_n = w$  и  $v_{i+1}$  для любого  $i$  ( $1 \leq i \leq n$ ) непосредственно выводимо из  $v_i$ . Последовательность цепочек  $v_0, v_1, \dots, v_n$ , где  $v_0$  есть символ «Предложение», каждая цепочка  $v_{i+1}$  непосредственно выводима из цепочки  $v_i$ , а цепочка  $v_n$  есть цепочка над словарем  $V$ , называется выводом.

**0. 3. 10.** Фразой, порождаемой грамматикой  $G$ , называется заключительная цепочка вывода.

Пусть  $f$  есть фраза, порождаемая грамматикой, а  $A \rightarrow B$  есть правило грамматики, участвующее в порождении фразы  $f$ . Тогда цепочку  $f'$  над словарем  $V$ , выводимую из символа  $A$ , мы будем называть составляющей фразы  $f$ , маркированной символом  $A$ .

В частности по этому определению сама фраза  $f$  является составляющей фразы  $f$ . Составляющая  $x$  называется непосредственно составляющей для составляющей  $y$ , если а)  $x$  и  $y$  суть разные составляющие одной и той же фразы  $f$ , б) символ  $X$ , маркирующий составляющую  $x$ , входит в цепочку, непосредственно выводимую из символа  $Y$ , который маркирует составляющую  $y$ . Совокупность составляющих фразы  $f$  мы будем называть системой составляющих фразы  $f$ .

**0. 3. 11.** Грамматикой непосредственно составляющих (или НС-грамматикой) называется грамматика, в которой каждое правило имеет вид  $\chi_1 y \chi_2 \rightarrow \chi_1 w \chi_2$ , где  $y$  элемент из  $V_1$ ,  $w$  непустая цепочка над  $V \cup V_1$ .

Иначе говоря, применение каждого правила НС-грамматики позволяет переписать  $y$  как  $w$  в окружении  $(\chi_1, \chi_2)$ .

**0. 3. 12.** Грамматика непосредственно составляющих, в которых  $\chi_1$  и  $\chi_2$  из определения 3. 11 суть пустые символы, т. е. каждое правило имеет вид  $y \rightarrow w$ , где  $y \in V_1$ , а  $w$  непустая цепочка над  $V \cup V_1$  называется контекстно свободной, или бесконтекстной. В противном случае грамматика непосредственно составляющих называется контекстно-связанной.

**0. 3. 13.** Грамматика непосредственно составляющих, в которой все правила из  $S$  имеют вид:  $A \rightarrow b$ ,  $A \rightarrow B$ ,  $A \rightarrow \Lambda$  или  $A \rightarrow aB$ , где  $A, B$  суть элементы из  $V_1$ ,  $\Lambda$  — пустая цепочка, а  $a$  и  $b$  суть элементы из  $V$ , называется грамматикой с конечным числом состояний, или автоматной грамматикой.

Из определения следует, что всякая грамматика с конечным числом состояний есть бесконтекстная грамматика.

Мы в дальнейшем ограничимся лишь бесконтекстными грамматиками. Правила, входящие в  $S$ , классифицируются следующим образом.

0. 3. 14. Пусть имеется некоторое правило  $A \rightarrow B$ , где  $A$  есть элемент  $V_1$ . Оно называется правилом воплощения (или правилом реализации), если  $B$  есть элемент  $V$ , правилом переименования, если  $B$  есть элемент  $V_1$  и правилом развертывания, если  $B$  содержит более одного элемента.

Практически все правила воплощения суть соответствия, обратные функции  $\Pi$ , сопоставляющей каждому слову его категориальный пучок, т. е. правила воплощения имеют вид:

категориальный пучок  $\rightarrow$  некоторое слово, имеющее данный категориальный пучок

Теперь мы дадим новое определение фразы (вообще говоря, это понятие отлично от того, которое определялось процедурой § 19, но поскольку практически оба понятия совпадают, мы будем для удобства пользоваться тем же термином, ибо в случае необходимости всегда можно говорить о правильности в смысле 3. 7 и правильности в смысле 3. 16).

0. 3. 15. Грамматика  $G$  совместима с языком  $L$ , если она производит все реальные фразы языка и не производит ни одной запрещенной цепочки.

0. 3. 16. Цепочка  $f$  есть фраза в языке  $L$ , если она производится грамматикой, совместимой с этим языком.

Ясно, что и здесь, как и в § 18, не являющаяся фразой цепочка необязательно должна быть запрещенной, но фраза уже не может совпадать с запрещенной цепочкой.

Приведем здесь очень простой пример грамматики:

$V = \{\text{девочка, идея, мальчик, спит, сказал, взглянуть, спокойно, равнодушно, яростно, очень, красивая, счастливая}\}$

$V_1 = \{\text{предложение, (существительное И, ед. ч. Ж), (существительное И, ед. М), (глагол 3-е л., ед. ч.), (инфинитив), (наречие образа действия), (наречие степени), (прилагательное И, ед. ч. Ж, положит.)}\}$

$S$ : а) правила развертывания:

предложение  $\rightarrow$  существительное И, ед. ч. Ж + глагол 3-е л. ед. ч.;

предложение  $\rightarrow$  существительное И, ед. ч. М + глагол 3-е л. ед. ч.;

предложение  $\rightarrow$  наречие образа действия + инфинитив;  
существительное И., ед. ч. Ж  $\rightarrow$  прилагательное И, ед. ч. Ж,  
положит. + существительное И, ед. ч. Ж.;

глагол 3-е л. ед. ч.  $\rightarrow$  наречие образа действия + глагол 3-е л. ед. ч.;

прилагательное И, ед. ч. Ж, положит. → наречие степени + прилагательное И, ед. ч. Ж, положит.

б) правила воплощения:

существительное И, ед. ч. Ж → *девочка, идея*;

глагол 3-е л., ед. ч. → *спит, говорит*;

прилагательное М, ед. ч. Ж, положит. → *красивая, счастливая*;

наречие образа действия → *спокойно, равнодушно, яростно*;

наречие степени → *очень*.

Эта грамматика, как легко видеть, совместима с языком, имеющим то же множество реальных фраз и запрещенных цепочек, что и в § 19.

В частности, фразой, порождаемой этой грамматикой, является фраза *очень красивая девочка спокойно спит*

Эта фраза имеет следующие составляющие:

1) *очень красивая девочка спокойно спит* — маркирована символом «предложение».

2) *очень красивая девочка* — маркирована символом «существительное И, ед. ч. Ж».

3) *красивая девочка* — маркирована символом «существительное И, ед. ч. Ж».

4) *девочка* — маркирована символом «существительное И, ед. ч. Ж.»

5) *очень красивая* — маркирована символом «Прилагательное И, ед. ч. Ж. положит.».

6) *красивая* — маркирована тем же символом.

7) *очень* — маркирована символом «наречие образа действия».

8) *спокойно спит* — маркирована символом «глагол 3-е л., ед. ч.»

9) *спит* — маркирована тем же символом.

Здесь сразу же видно, что мы пользуемся системой обозначений, в которой словосочетание и его опорное слово маркируются одним и тем же символом. Можно, разумеется, ввести систему обозначений, в которой этого не происходит, но нам важно было подчеркнуть роль категорий опорного слова в процессе порождения.

Ясно, что указание категорий необходимо для того, чтобы не была произведена, например, запрещенная цепочка \* *счастливая мальчик*. Между тем указания на часть речи избыточны, ибо каждая часть речи однозначно определяется своим категориальным пучком. Это и дало нам возможность считать, что все правила воплощения могут строиться как соответствие между категориальным пучком и входящими в него словами. Ясно, что при расширении V и соответственно правил воплощения за счет других слов русского языка, входящих в те же категориальные пучки, что и приведенные слова, мы не получим запрещенных цепочек, вернее не получим ни одной цепочки, не являющейся правильно построенной фразой русского языка.

Таким образом, при наличии подобной грамматики и при данной интерпретации правил замена слова другим словом, входящим в тот же категориальный пучок, не меняет правильности цепочек.

## § 21. РАЗБИЕНИЯ СЛОВАРЯ И ОБРАЗЫ ЦЕПОЧЕК

Пусть нам дано произвольное разбиение  $V$  словаря  $V$  на непересекающиеся классы  $B_1, B_2, \dots, B_n$ . Можно привести следующие примеры разбиений:

1. **Единичное разбиение ( $E$ )**, где каждое подмножество  $B_i$  состоит ровно из одного слова. Мы будем говорить в этом случае о  $E$ -разбиении.

2. **Алфавитное разбиение  $\alpha$** , где каждое подмножество  $B_i$  состоит из тех и только тех слов, которые начинаются с данной буквы алфавита. Мы будем говорить в этом случае о  $\alpha$ -разбиении.

3. **Обратно-алфавитное разбиение  $\omega$** , где каждое подмножество  $B_i$  состоит из тех слов, которые оканчиваются на данную букву.

4. **Разбиение на парадигмы**, которое мы будем называть  $\Gamma$ -разбиением.

5. **Разбиение на категориальные пучки  $\Pi$** .

0. 3. 17.  $B$ -образом цепочки  $f = x_1 \dots x_n$  называется последовательность классов разбиения  $B$ , каждый из которых содержит соответствующее слово  $x_i$ .  $B$ -образ цепочки  $f$  обозначается через  $B(f)$ .

0. 3. 18.  $B$ -образ называется **отмеченным**, если существует хотя бы одна фраза, имеющая данный  $B$ -образ.

0. 3. 19.  $B$ -образ называется **совершенным**, если любая цепочка, имеющая данный  $B$ -образ, есть фраза.

Приведем теперь примеры  $B$ -образов. Возьмем фразу

*Очень маленькая девочка спит.*

$E$ -образом данной фразы будет последовательность классов

{очень} {маленькая} {девочка} {спит},

где каждый класс содержит ровно одно слово, поэтому нет никакой другой цепочки, имеющей данный  $B$ -образ.

$\alpha$ -образом данной фразы будет последовательность классов

О М Д С

Тот же  $\alpha$ -образ будет иметь, например, цепочка

*оленья микрористые душа с*

ω-образом данной фразы будет последовательность классов:

-ь -я -а -т

Тот же ω-образ имеет, например, цепочка

*конь красивая дама летит*

Г-образом данной фразы будет последовательность

*Г (очень) Г (маленькая) Г (девочка) Г (спит)*

Тот же Г-образ будет иметь цепочка

*очень маленькой девочками спать*

Приведем наконец II-образ данной фразы: (наречие степени) (И, Ж, ед. ч. положит.) (И, Ж, ед. ч.). (3-е л., ед. ч. наст.)

Тот же II-образ имеют фразы

*очень красивая дама летит*

*весьма приятная зебра гуляет* и т. п.

При разбиении *E* любой *E*-образ либо неотмечен, либо совершенен.

При α-разбиении α-образ любой цепочки является отмеченным, ибо практически всегда можно найти фразу, имеющую искомым α-образ. Мы видели, что всегда можно найти цепочку, не являющуюся фразой, но имеющую тот же α-образ.

При разбиении ω имеется, по-видимому, довольно много совершенных ω-образов: таково положение в тех случаях, когда в соответствующих фразах употреблены флективные неомонимичные формы, однако приведенный выше пример показывает, что ω-образы, вообще говоря, несовершенны.

Наконец, совершенными являются II-образы, ибо любая цепочка, имеющая тот же II-образ, что и некоторая фраза, сама является фразой (это во всяком случае верно при той интерпретации грамматики, которая дана в конце § 20).

**П р и м е ч а н и е.** В «Моделях языка» так же, как и в работе О. С. Кулагиной, впервые использовавшей соответствующее понятие, вместо термина «*B*-образ» использовался термин «*B*-структура».

Вообще говоря, такое словоупотребление не противоречит определению термина «структура» в данной книге (ср. определение 2. 27), ибо в качестве отношения, в связи с которым устанавливается изоморфизм, можно взять отношение линейной упорядоченности<sup>44</sup>. Однако в дальнейшем, говоря о графе фразы, мы будем иметь в виду отношение синтаксического подчинения и поэтому термин «структура фразы» будет употребляться в другом смысле. Исходя из этих соображений, мы предпочли отказаться от нашего

<sup>44</sup> Любопытно, что определение 2. 27 близко к описанию «структуры», даваемому Расселом (Б. Р а с с е л. Человеческое познание. М., 1957, стр. 288—290), и, с другой стороны, под «структурой предложения» Рассел понимает именно то, что мы определяем как *B*-образ фразы.

употребления термина «структура» и ввести термин « $B$ -образ». (Это не противоречит определению 2. 25, ибо соответствие между цепочкой и  $B$ -образом есть отображение в строгом смысле слова).

## § 22. ПОНЯТИЯ КОНТЕКСТА И ЗАМЕЩЕНИЯ. СЕМЕЙСТВО

0. 3. 20. Назовем контекстом<sup>45</sup> любую упорядоченную пару цепочек  $(f, g)$ , где цепочки  $f$  и  $g$  не обязательно фразы и могут быть и пустыми.

Примеры контекстов:

$(Я, рано)$ ,  $(в, \Lambda)$ ,  $(кого, видел)$ ,  $(кого, что)$   $(сегодня мы, рано)$   $(\Lambda, \Lambda)$  и т. п. (Напомним, что через  $\Lambda$  мы обозначаем пустую цепочку).

0. 3. 21. Мы будем говорить, что слово  $x$  (или цепочка  $X$ ) до п у с к а е т с я данным контекстом  $(f, g)$ , если цепочка  $fxg$  (или соответственно  $fXg$ ) есть фраза.

0. 3. 22. Мы будем говорить, что парадигма  $\Gamma(x)$  до п у с к а е т с я контекстом  $(f, g)$ , если существует  $x' \in \Gamma(x)$ , такое, что  $(f, g)$  допускает  $x'$ .

Аналогично определим понятия произвольного  $B$ -контекста.

0. 3. 23. Назовем  $B$ -контекстом любую упорядоченную пару  $B$ -образов.

0. 3. 24. Множество  $B_i$  до п у с к а е т с я данным  $B$ -контекстом  $(B(f), B(g))$ , если  $B$ -образ  $B(f)$   $B_i$   $B(g)$  отмечен.

0. 3. 25. Мы будем говорить, что слово  $x$  з а м е щ а е т  $y$  (записывается  $y \Rightarrow x$ ), если любой контекст, допускающий  $y$ , допускает и  $x$  (мы будем употреблять выражения « $x$  замещает  $y$ » и « $y$  замещается  $x$ » как равнозначные).

0. 3. 26. Мы будем говорить, что  $x$  в з а и м о з а м е щ а е м о с  $y$ , или  $x$   $E$ -э к в и в а л е н т н о  $y$ , если имеет место  $x \Rightarrow y$  и  $y \Rightarrow x$ .

Понятие замещения обобщает идею местоименности, играющую столь важную роль в грамматике. Что значит, что слово *он* является местоимением (мужского рода именительного падежа, единственного числа)? Мы можем в наших терминах выразить это следующим образом. Возьмем какое-нибудь существительное ( $M, И, ед. ч.$ ), например *слон*. В любом контексте, в котором употребляется слово *он*, можно заменить *он* на *слон*, в то время, как обратное неверно, ибо, например, в контексте (*этот, сердит*), нельзя заменить *слон* на *он*.

<sup>45</sup> Термин «контекст» в этом смысле предложен С. Маркусом (см.: S. Marcus. Modèles mathématiques pour la catégorie grammaticale du cas. «Revue roumaine de mathématiques pures et appliquées», 1963, № 4). Соответствующее понятие рассматривалось Р. Л. Добрушиным, который употреблял термин «фраза с многоточием» (см.: Р. Л. Добрушин. Элементарная грамматическая категория. «Бюллетень объединения по проблемам машинного перевода», № 5, 1957, стр. 19—20).



В нашем определении мы и отразили эту идею, говоря, что слово *он* замещает слово *слон*. Однако принятое нами формальное определение гораздо шире и распространяется на ряд других явлений (часть из которых мы рассмотрим в § 25).

**Примечание.** В «Моделях языка» и ряде других работ вместо «*x* замещает *y*» говорится «*y* подчиняет *x*». Однако термин «подчинение», во-первых, является слишком общим, так как применим для любого иерархического отношения, а, во-вторых, он получил широкое распространение в синтаксических работах, связанных с лингвистическим моделированием. Поэтому здесь мы пользуемся термином «замещение», апеллирующим к лингвистической интуиции<sup>46</sup>. Что касается термина «взаимозамещаемость», то мы вводим этот новый термин наряду со старым термином «эквивалентность», ибо «эквивалентность слов» ни в каком другом смысле не обсуждается.

Обобщим теперь идею замещения на произвольные *B*-разбиения.

**0. 3. 27.** Мы будем говорить, что класс  $B_j$  замещает класс  $B_i$  (записывается  $B_i \xrightarrow{B} B_j$ ), если любой *B*-контекст, допускающий  $B_j$ , допускает и  $B_i$ .

**0. 3. 28.**  $B_i$  *B*-эквивалентно  $B_j$  (записывается  $B_i \sim_B B_j$ ), если  $B_i \xrightarrow{B} B_j$  и  $B_j \xrightarrow{B} B_i$ . В частном случае, если  $x \Rightarrow y$  и  $y \Rightarrow x$ , то  $x \sim_E y$ .

*B*-эквивалентность обладает, как легко проверить, свойствами симметрии, рефлексивности и транзитивности.

**0. 3. 29.** П р о и з в о д н ы м разбиением  $B^*$  от разбиения  $B$  называется разбиение, при котором в одно множество  $B^*(x)$  попадают такие и только такие множества  $B(y)$ , что  $B(y) \sim_B B(x)$ .

**0. 3. 30.** Разбиение, производное от *E*-разбиения, называется разбиением на с е м е й с т в а.

Таким образом, в семейство  $S(x)$  попадают все  $y$ , такие, что  $y \Rightarrow x$  и  $x \Rightarrow y$ . Приведем еще один пример производного разбиения, а именно все множество слов, которое является производным от  $\alpha$ -разбиения (это следует из сделанного выше замечания, что любой  $\alpha$ -образ в славянских языках, в частности в русском языке, отмечен). Другие примеры даются в § 26.

Особенно важными для дальнейшего изложения будут понятия *E*-эквивалентности и семейства. Прежде всего отметим важное свойство *S*-образов (т. е. образов, где в качестве множеств  $B_i$  выступают семейства).

**Т. 3. 2.** а) Всякий отмеченный *S*-образ совершенен и б) если при некотором разбиении *B* всякий *B*-образ или не отмечен или же совершенен, то для любого  $x$  верно  $\bar{B}(x) \subseteq S(x)$ .

<sup>46</sup> Термин «замещение» введен, сколько известно автору, А. В. Гладким (см.: А. В. Г л а д к и й. О распознаваемости замещаемости в рекурсивных языках. «Алгебра и логика. Семинар», 1963; 2—3; стр. 5—22).

Выясним соотношение между отношением  $E$ -эквивалентности (взаимозамещаемости) и рассмотренными в § 19—20 способами порождения множества фраз.

Прежде всего обратимся к связи между понятиями дистрибутивной близости и  $E$ -эквивалентности (взаимозаменяемости).

Покажем, что два дистрибутивно близких слова  $E$ -эквивалентны, но обратное, вообще говоря, неверно.

В самом деле, пусть  $x$  дистрибутивно близко с  $y$ . Напомним, что множество фраз по определению 3. 7 формируется таким образом, что если цепочка  $fxg$  была фразой и  $x$  дистрибутивно близко с  $y$ , то цепочка  $fyg$  будет также фразой. Это будет иметь место для произвольных контекстов  $(f, g)$ . А отсюда сразу следует, что  $x \in S(y)$ . Обратное, разумеется, не верно. Пусть нам даны реальные фразы:

1. Он чутко спал
2. Он спокойно сидел
3. Я сидел
4. Я спал

Тогда *сидел* дистрибутивно близко со *спал* и фразами будут цепочки:

5. Он спокойно спал
6. Он чутко сидел

Других фраз нет. Это означает, что на данном множестве фраз *спокойно* и *чутко* входят в одно семейство, но не находятся в отношении дистрибутивной близости. Достаточно, однако, ввести одну из фраз (5 или 6) в число реальных, чтобы отношения  $E$ -эквивалентности и дистрибутивной близости совпали. Таким образом, всегда существует такое «разумное» задание множества реальных фраз, при котором оба отношения совпадают.

Здесь необходимо сделать одно разъяснение (которым автор обязан А. В. Гладкому). Совпадение отношений  $E$ -эквивалентности и дистрибутивной близости возможно лишь в пределе. Дело в том, что для любого конечного множества фраз можно так доопределить систему реальных фраз, чтобы на этом множестве дистрибутивная близость совпала с  $E$ -эквивалентностью. Неясно, можно ли это сразу сделать для всего языка.

И все же существует интуитивная уверенность в том, какие реальные фразы и какие запрещенные цепочки нужно задать для проверки каждой конкретной  $E$ -эквивалентности.

По-видимому, практически можно проверять именно дистрибутивную близость и отсюда делать заключения о  $E$ -эквивалентности (ниже мы рассмотрим условия, при которых можно ограничиваться отдельными фрагментами языка для такой проверки).

Пусть теперь данное множество фраз порождается другой процедурой, а именно — некоторой грамматикой по определению 3. 16. Мы договорились о такой интерпретации этой грамматики, при которой в правилах воплощения категориальный

пучок заменяется некоторым словом, входящим в него. Это означает, что если фраза  $fxg$  произведена данной грамматикой, а слово  $y$  относится к тому же категориальному пучку, что и  $x$ , то фраза  $fug$  также будет произведена той же грамматикой. Итак, если  $x$  относится к тому же категориальному пучку, что и  $y$ , то  $x$  относится к тому же семейству, что и  $y$ . Обратное, вообще говоря, неверно; два слова могут быть  $E$ -эквивалентными, но относиться к разным категориальным пучкам.

Однако и здесь опыт показал, что в естественных языках (в частности, славянских) положение таково: чем больше фактов данного естественного языка принимается во внимание, тем вероятнее совпадение категориального пучка с семейством.

Особенно показателен следующий пример. Во многих рассуждениях о русском языке в терминах теоретико-множественных понятий существенную роль играло утверждение, что в русском языке все слова одного падежа во множественном числе  $E$ -эквивалентны, т. е. образуют одно семейство. Так, утверждалось, что слово *столы* всегда можно заменить на слово *лампы*, а также на слово *коровы* без изменения грамматической правильности.

Однако если включить в число запрещенных следующие цепочки:

\*Я вижу столы, одну из которых я купил

\*Я вижу лампы, один из которых поломан

то замена *коровы* на *столы* или *столы* на *лампы* становится невозможной<sup>47</sup>. Итак, на всем множестве эти слова, относящиеся к разным категориальным пучкам, входят и в разные семейства.

Все эти соображения позволяют сформулировать гипотезу о том, что на всем множестве фраз семейство совпадает с категориальным пучком. Эту гипотезу мы условно назовем основным тезисом лингвистической теории моделей.

0. 3. 31. Мы будем говорить, что контекст  $(f_1, g_1)$  является более емким, чем контекст  $(f_2, g_2)$ , если любое слово, допускаемое контекстом  $(f_1, g_1)$ , допускается контекстом  $(f_2, g_2)$ . Тот факт, что  $(f_1, g_1)$  более емко, чем  $(f_2, g_2)$ , мы будем записывать следующим образом

$$(f_1, g_1) \Rightarrow (f_2, g_2)$$

Примечание. Мы пользуемся той же записью для большей емкости, что и для отношения замещения, ибо с формальной точки зрения это близкие понятия, а, с другой стороны, невозможно смешать эти два отношения, так что эта запись не поведет к кривотолкам.

0. 3. 32. Если  $(f_1, g_1) \Rightarrow (f_2, g_2)$  и  $(f_2, g_2) \Rightarrow (f_1, g_1)$ , то мы говорим, что контексты  $(f_1, g_1)$  и  $(f_2, g_2)$  равноемки и записываем это в виде  $(f_1, g_1) \sim (f_2, g_2)$ .

<sup>47</sup> Это тонкое наблюдение сделано А. А. Зализняком.

Равноемкость контекстов, так же как и  $E$ -эквивалентность слов, обладает свойствами рефлексивности, симметрии и транзитивности, т. е. есть отношение типа эквивалентности и поэтому все множество контекстов разбивается на непересекающиеся подмножества равноемких контекстов.

**0. 3. 33.** Множество всех равноемких контекстов называется контекстным классом.

**0. 3. 34.** Контекстный класс  $K_1$  является более емким, чем контекстный класс  $K_2$  (записывается  $K_1 \Rightarrow K_2$ ), если имеются  $(f_1, g_1) \in K_1$  и  $(f_2, g_2) \in K_2$ , такие, что  $(f_1, g_1) \Rightarrow (f_2, g_2)$ .

Вообще говоря, данным определением еще нельзя пользоваться, ибо логически мыслима ситуация, когда некий контекст из  $K_1$  более емок, чем какой-то контекст из  $K_2$ , но зато в  $K_2$  имеется другой контекст, который более емок, чем некоторый контекст в  $K_1$ . Следующее утверждение показывает, что определение вполне корректно.

**Т. 3. 3.** Если  $K_1 \Rightarrow K_2$ , то любой контекст, входящий в  $K_1$ , является более емким, чем любой контекст, входящий в  $K_2$ .

Понятия емкости контекстов и контекстных классов понадобятся нам при моделировании понятия «падеж» в § 40.

## § 23. ПОНЯТИЕ ПРАВИЛЬНОГО ФРАГМЕНТА

До сих пор мы рассматривали  $E$ -эквивалентность относительно всего множества фраз  $\theta$ . Между тем фактически проверка  $E$ -эквивалентности (сводящаяся, как мы видели, к проверке дистрибутивной близости) может производиться лишь на некоторых частях множества фраз. Это вполне соответствует лингвистической практике, ибо лингвист всегда имеет дело с некоторой частью всех фраз языка.

С другой стороны, очевидно, что фрагменты, выбираемые нами, должны в какой-то мере отражать существенные свойства языка, т. е. не быть случайными выборками фраз.

Эту задачу можно решить следующим образом. Пусть нам известно разбиение на семейства в исходном множестве фраз  $\theta$ . В силу основного тезиса теории моделей для этого достаточно, чтобы была задана функция  $\Pi$ , ставящая в соответствие каждому слову его категориальный пучок. Рассмотрим некоторое подмножество  $\Delta$  всего множества фраз  $\theta$ . Для  $\Delta$  тем способом, что и для  $\theta$ , можно определить отношения подчинения и  $E$ -эквивалентности. Мы будем записывать при помощи  $x \Rightarrow_{\Delta} y$  тот факт, что  $x$  замещается  $y$  относительно фрагмента  $\Delta$ . Аналогично  $x \sim_{\Delta} y$  обозначает, что  $x$   $E$ -эквивалентно  $y$  относительно фрагмента  $\Delta$ . При этом  $x$  и  $y$  входят в одно семейство в том и только в том случае, если  $x \sim_{\theta} y$ , где  $\theta$  — все множество фраз.

0. 3. 35. Фрагмент  $\Delta$  называется правильным, если для любых  $x$  и  $y$  из  $x \underset{\theta}{\sim} y$  следует  $x \underset{\Delta}{\sim} y$ .

Заметим, что все множество фраз есть правильный фрагмент (так же как и пустое множество). Почти очевидны также следующие утверждения.

Т. 3. 4. а) Дополнение правильного фрагмента  $\Delta$  до всего множества фраз, т. е.  $\bar{\Delta}$  есть правильный фрагмент; б) объединение двух правильных фрагментов есть правильный фрагмент; в) пересечение двух правильных фрагментов есть правильный фрагмент.

Итак фактически проверка  $E$ -эквивалентности (сводящаяся к проверке дистрибутивной близости) должна производиться лишь на правильных фрагментах, причем, исходя из теоремы 3. 4, легко из одних правильных фрагментов строить другие. Все это приближает теорию моделей к лингвистической практике.

Приведем еще несколько примеров правильных фрагментов:

Т. 3. 5 а) Для любого  $k$  множество фраз длины, не превышающей  $k$ , есть правильный фрагмент.

б) Для любого  $k$  множество всех данных фраз, в которых не содержится слов из  $k$  зафиксированных семейств, есть правильный фрагмент.

Другие примеры мы приведем далее.

Нам понадобится в дальнейшем следующее видоизменение отношения замещения:

0. 3. 36. Мы будем говорить, что  $x$  ограничено замещается  $y$ , если существует хотя бы один правильный фрагмент  $\Delta$  такой, что  $x \underset{\Delta}{\Rightarrow} y$ , и  $x$  входит хотя бы в одну фразу из  $\Delta$ .

## § 24. ЯЗЫК И ПОДЪЯЗЫК

В предыдущих параграфах (точнее в § 14—19) мы рассмотрели основные понятия, необходимые для построения лингвистических моделей. Теперь дадим определение языка в терминах теории моделей:

0. 3. 37. Языком  $L$  называется совокупность следующих объектов:

- а) некоторое множество слов  $V = \{x\}$ ;
- б) выделенное множество реальных фраз  $\alpha$ , причем каждое слово из  $V$  входит хотя бы в одну реальную фразу;
- в) выделенное множество запрещенных цепочек  $\beta$ ;
- г) разбиение множества  $V$  на парадигмы  $\Gamma$ ;
- д) система  $\kappa$ , подмножеств множества  $V$ , называемых категориями, причем для каждого слова известна функция  $\Pi$ , ставящая в соответствие слову его категориальный пучок.

Сокращенно определение 3. 37 можно записать формулой

$$L = (V, \alpha, \beta, \Gamma, \kappa, \Pi).$$

Поскольку задание множества реальных и запрещенных цепочек однозначно определяет некоторое множество фраз  $\theta$ , то язык можно (а для дальнейшего и более удобно) рассматривать в форме

$$L = (V, \theta, \Gamma, \kappa, \Pi).$$

Язык  $L$  называется категориально пустым, если  $\kappa$  и  $\Pi$  не заданы. Язык  $L$  называется полностью интерпретированным, если на множестве  $\alpha$  задано отношение смыслового включения (ср. § 18).

**0.3.38.** Подязыком  $L'$  данного языка  $L$  называется совокупность объектов  $L' = (V', \Delta, \Gamma', \kappa', \Pi)$ , где  $V' \subseteq V$ ,  $\Delta$  есть правильный фрагмент всего множества  $\theta$ , причем каждое слово из  $V'$  входит в хотя бы одну фразу из  $\Delta$ ,  $\Gamma'$  есть такое разбиение на парадигмы, что для любых слов  $x$  и  $y$  (из  $V$ ) верно следующее: если  $x \in \Gamma'(y)$ , то  $x \in \Gamma(y)$ , а  $\kappa'$  система подмножеств множества  $V'$ , таких, что любое  $\kappa'$  есть подмножество некоторой категории  $\kappa$ .

## § 25. ЭЛЕМЕНТАРНАЯ ГРАММАТИЧЕСКАЯ КАТЕГОРИЯ И ПРОБЛЕМА ОМОНИМИИ

С помощью понятия замещения был сформулирован чрезвычайно важный понятийный аппарат, первоначальная идея которого изложена Р. Л. Добрушиным<sup>48</sup>. Мы изложим его однако в несколько более общей форме, исходя отчасти из некоторых работ С. Маркуса<sup>49</sup>.

**3.39.** Пусть у нас имеется произвольное разбиение  $B$  и некоторый класс этого разбиения  $B(x)$  и пусть не существует множества  $B(y)$  такого, что  $B(y) \neq B(x)$  и  $B(y) \Rightarrow_B B(x)$ .

Тогда класс  $B(x)$  называется начальным.

**Т.3.6.** Если  $B(x) \Rightarrow_B B(y)$ , то  $B^*(x) \Rightarrow_B B^*(y)$

Следствие: Если  $B(x)$  начальный класс, то  $B^*(x)$  также начальный класс.

Рассмотрим в качестве примера единичное разбиение. По теореме 3.6 из  $x \Rightarrow y$  следует  $S(x) \Rightarrow S(y)$ .

**0.3.40.** Назовем рядом, порожденным множеством  $B(x)$ , совокупность  $B(x) \cup B(y_1) \cup \dots \cup B(y_n)$ , где 1) для любого  $i$  ( $1 \leq i \leq n$ ) имеет место  $B(x) \Rightarrow_B B(y_i)$ , 2) если существует класс  $B(z)$  такой, что  $B(x) \Rightarrow_B B(z)$ , то найдется  $i$ , такое, что  $B(y_i) = B(z)$ .

<sup>48</sup> Р. Л. Добрушин. Элементарная грамматическая категория. «Бюллетень объединения по проблемам машинного перевода». М., 1958, стр. 19—20.

<sup>49</sup> S. Marcus. *Lingvistică matematică*. București, 1963.

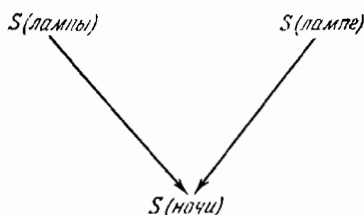
0. 3. 41. Ряд, порожденный некоторым начальным семейством  $S(x)$ , называется элементарной грамматической категорией. Таким образом, всякая элементарная грамматическая категория есть объединение семейств. Понятие элементарной грамматической категории позволяет поставить вопрос о грамматической омонимии.

0. 3. 42. Слово  $x$  называется омонимичным, если оно входит в две разные элементарные грамматические категории.

0. 3. 43. Слово  $x$  называется омонимом первого рода, если оно омонимично, но во всех начальных семействах  $S_1, \dots, S_k$ , породивших элементарные категории, в которые входит слово  $x$ , найдутся слова  $y_1 \in S_1, \dots, y_k \in S_k$ , входящие в одну и ту же парадигму.

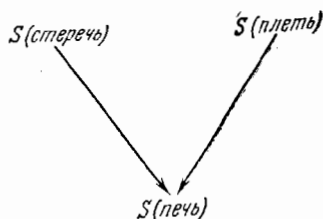
0. 3. 44. Слово  $x$  называется омонимом второго рода, если оно омонимично и не является омонимом первого рода.

Лингвистическую значимость предложенного разграничения омонимов иллюстрируют следующие примеры. Имеют, например, место отношения:



Здесь слово *ночи* входит в две элементарные категории, но является омонимом первого рода, так как во всех начальных семействах найдутся слова из одной парадигмы.

Рассмотрим теперь отношения:



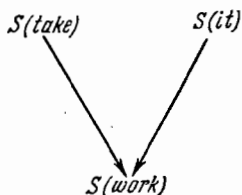
Здесь слово *печь* также входит в две элементарные грамматические категории, однако нет такой парадигмы, которая пересекалась бы со всеми начальными семействами, породившими эти категории. Это омоним второго рода.

Из этих примеров очевидно, что омонимы первого рода чрезвычайно распространены в флективных языках, в особенности

славянских, явление, обычно называемое п е й т р а л и з а ц и е й. Отвлечься от этого столь характерного для славянских языков явления нельзя даже в модели, если мы хотим, чтобы модель отражала существенные черты исследуемых языков.

Иначе обстоит дело с омонимией второго рода, приводящей к совпадению разных частей речи — существительного и глагола.

Для славянских языков такие случаи крайне редки, они носят случайный характер и от них легко можно отвлечься. С другой стороны, в английском языке такая омонимия чрезвычайно распространена, ср.



где *work* и масса аналогичных слов одновременно существительные и глаголы.

Напомним, что мы договорились в § 15 заранее разграничить такие омонимы. Можно представить себе дело таким образом, что омонимы второго рода получают в контекстах, где они взаимозаменяемы со словами из одного начального семейства, один индекс, а в контекстах, где они взаимозаменяемы со словами другого начального семейства — другой индекс, например, слово *печь* в контексте (*я вижу, Λ*) получает особый индекс, отличный от индекса, который получает слово *печь* в контексте (*начал, пироги*). Иначе говоря, слово *печь* распадается на два слова *печь*<sub>1</sub> и *печь*<sub>2</sub>.

В подобных случаях мы будем говорить о «расщеплении» омонимов. Понятие элементарной грамматической категории можно обобщить следующим образом.

Мы знаем, что отношение  $x \Rightarrow y$  является отношением рефлексивным, транзитивным и несимметричным. Его можно превратить в симметричное разными способами. Один из способов мы уже рассмотрели: два слова называются *E*-эквивалентными, если одновременно имеет место  $x \Rightarrow y$  и  $y \Rightarrow x$ . Рассмотрим теперь аналогическое отношение, где связка «и» заменяется связкой «или».

**0. 3. 45.** Два слова  $x$  и  $y$  непосредственно связаны *а*, если выполняется хотя бы одно из условий  $x \Rightarrow y$  или  $y \Rightarrow x$ . Отношение непосредственной связанности рефлексивно симметрично, но, вообще говоря, нетранзитивно. Возьмем однако транзитивное замыкание этого отношения.

**0. 3. 46.** Два слова  $x$  и  $y$  относятся к одному классу связанности, или *D*-классу, если их можно соединить це-



почкой  $z_1 = x_1$ ,  $z_n = y$ , в которой для любого  $i$  верно:  $z_i$  и  $z_{i-1}$  находятся в отношении непосредственной связи.

Таким образом, каждый  $D$ -класс представляет собой связный подграф в графе отношений замещения.

Надо сказать, что, если бы мы не расщепили омонимов второго рода, то  $D$ -классы не представляли бы собой лингвистической ценности, ибо в один  $D$ -класс попадали бы слова разных частей речи.

Теперь же  $D$ -класс представляет собой объединение семейств внутри некоторой части речи, причем в один  $D$ -класс два семейства попадают лишь в случае наличия некоторых нейтрализаций, например семейства  $S$  (лампы) и  $S$  (лампе) попадают в один  $D$ -класс потому, что в семействе  $S$  (ночи), также входящем в этот  $D$ -класс, нейтрализуются формы двух падежей.

## МОДЕЛИРОВАНИЕ ЧАСТИ РЕЧИ

## § 26. ПОНЯТИЕ ТИПА

Исследование «теоретико-множественными» методами наиболее абстрактных свойств языков началось с постановки вопроса о моделировании «части речи». Закономерно при этом, что такое центральное понятие традиционной лингвистической теории, как понятие «части речи», оказалось весьма плодотворным исходным пунктом для создания большого раздела теории моделей, хотя лингвистическая ценность этого понятия и не сразу была осознана<sup>1</sup>.

Определение первого аналога части речи, а именно типа, состоит в следующем.

**0.4.1.** Два слова  $x$  и  $y$  относятся к одному типу, если для любого контекста  $(z_1 \dots z_k, z_{k+1} \dots z_n)$ , допускающего  $\Gamma(x)$ , найдется контекст  $(z'_1 \dots z'_k, z'_{k+1} \dots z'_n)$ , допускающий  $\Gamma(y)$  и такой, что для любого  $i$  ( $1 \leq i \leq n$ )  $z'_i \in \Gamma(z_i)$ , и обратно, для любого контекста  $(u_1 \dots u_l, u_{l+1} \dots u_m)$ , допускающего  $\Gamma(y)$ , найдется контекст  $(u'_1 \dots u'_l, u'_{l+1} \dots u'_m)$ , допускающий  $\Gamma(x)$  и такой, что для любого  $j$  ( $1 \leq j \leq m$ )  $u'_j \in \Gamma(u_j)$ .

Покажем, что по этому определению слова *мальчик*, *кукла*, *сито* входят в один тип. Возьмем произвольный контекст, допускающий слово *мальчик*, например (*маленький, лежал на полу*), но для этого контекста можно найти контексты (*маленькая, лежала на полу*), (*маленькое, лежало на полу*), допускающие соответственно слова *девочка* и *кукла*, причем в этих контекстах все слова на соответственных местах принадлежат одной парадигме.

Подобным же образом можно было бы показать, что в один тип объединяются такие прилагательные, как *большой*, *красивая*, *сильное* и т. п., в другой тип объединяются такие глаголы, как *лежать*, *стоять*, *идти* и т. п.

С формальной точки зрения интересно, что разбиение на типы есть производное от разбиения на парадигмы, т. е.  $x$  и  $y$  попадают

<sup>1</sup> См. § 24, 27 в «Моделях языка», где говорится об «ограниченной ценности понятия части речи».

В один тип тогда и только тогда, когда  $\Gamma(x) \sim \Gamma(y)$ . Этот факт полезен для более наглядного представления, почему, например, приведенные прилагательные попадают в один тип. Возьмем такие парадигмы:

$\Gamma$  (большой) = {большой, большого, . . . , большом, большая, большое, большие, больших}

$\Gamma$  (красивая) = {красивый, красивого, . . . , красивом, красивая, красивое, красивые, красивых}

$\Gamma$  (сильное) = {сильный, сильного, . . . , сильном, сильная, сильное, сильные, сильных}.

Иначе говоря, мы возьмем все формы соответствующих прилагательных. Очевидно, что получающиеся парадигмы эквивалентны и, стало быть, входят в один тип.

Для существительных определение через производное разбиение несколько труднее (см.: «Модели языка», стр. 82—83), и для них приведенное нами определение представляется более удобным.

Очень интересно рассмотреть в связи с понятием типа вопрос о местоимениях-существительных, в первую очередь о личных местоимениях. Как известно, вопрос о распределении личных местоимений по парадигмам решается неоднозначно<sup>2</sup>, например, неясно, считать ли *я* и *мы*, *она*, *оно* и *они* — формами одного слова или разными словами. Бесспорно лишь то, что падежные формы, например, *я* и *меня* относятся к одной парадигме. Интересно поэтому, что какое бы разбиение на парадигмы мы ни взяли, все указанные местоимения будут принадлежать к одному типу, если падежные формы объединены в одной парадигме. Возьмем, например, разбиение, соответствующее, сколько можно судить, распределению по парадигмам, предлагаемому Академической грамматикой<sup>3</sup>.

$\Gamma_1$	$\Gamma_2$	$\Gamma_3$	$\Gamma_4$	$\Gamma_5$	$\Gamma_6$	$\Gamma_7$
<i>я</i>	<i>ты</i>	<i>он</i>	<i>мы</i>	<i>вы</i>	<i>кто</i>	<i>что</i>
<i>меня</i>	<i>тебя</i>	<i>его</i>	<i>нас</i>	<i>вам</i>	<i>кого</i>	<i>чего</i>
<i>мне</i>	<i>тебе</i>	<i>ему</i>	<i>нам</i>	<i>вами</i>	<i>кому</i>	<i>чему</i>
<i>мной</i>	<i>тобой</i>	<i>им</i>	<i>нами</i>	<i>вас</i>	<i>кем</i>	<i>чем</i>
		<i>она</i>				
		<i>ее</i>				
		<i>ей</i>				
		<i>оно</i>				
		<i>они</i>				
		<i>их</i>				
		<i>ими</i>				

<sup>2</sup> См.: В. В. Виноградов. Современный русский язык. М., 1938, стр. 159—160.

<sup>3</sup> «Грамматика русского языка», т. I, М. Изд-во АН СССР, 1952, стр. 388—390, 395.

Покажем, например, что *мной* и *вам* принадлежат одному типу. Возьмем произвольный контекст, допускающий некоторое слово из  $\Gamma_1$ , например:  $(\Lambda, \text{пришла})$  или  $(\text{доволен}, \Lambda)$ , но мы сразу же укажем контексты, допускающие слова из  $\Gamma_5$  и имеющие на соответственных местах слова из тех же парадигм:  $(\Lambda, \text{пришли})$ ,  $(\text{доволен}, \Lambda)$  аналогично для всех других существительных-местоимений. Единственное местоимение, образующее особый тип — это местоимение *себя*, поскольку в его парадигме нет слова, допускаемого одним из контекстов вида  $(\Lambda, \text{пришел})$ ,  $(\Lambda, \text{пришла})$ ,  $(\Lambda, \text{пришло})$ ,  $(\Lambda, \text{пришли})$ .

Приведенные примеры показывают, что понятие типа довольно близко к понятию «части речи» (более близкому аналогу будет посвящен следующий параграф).

Важным свойством отношения «принадлежать к одному типу» является то, что оно есть отношение эквивалентности, разбивающее все множество слов на непересекающиеся подмножества слов разных типов.

Рассмотрим теперь еще следующие понятия.

**0.4.2.** Два слова  $x$  и  $y$  связаны  $\Gamma$ - $S$ -цепью, если можно найти  $n$  слов  $z_1, \dots, z_n$  таких, что  $z_1 = x$ ,  $z_n = y$  и для любого  $i$  ( $1 \leq i < n$ ) верно  $z_i \in \Gamma(z_{i+1})$  или  $z_i \in S(z_{i+1})$ .

**0.4.3.** Разделом  $Ra(x)$  слова  $x$  называются совокупность всех слов  $y$ , связанных с  $x$  при помощи  $\Gamma$ - $S$ -цепей.

Отношение «принадлежать одному разделу» может быть получено и следующим образом. У нас имеются два отношения:  $\Gamma$  — «принадлежать одной парадигме» и  $S$  — «принадлежать одному семейству». Возьмем теперь новое отношение  $\Gamma \cup S$  и образуем транзитивное замыкание этого отношения  $(\Gamma \cup S)^+$ . Тогда мы и получим отношение «принадлежать одному разделу». Легко видеть, что это отношение есть отношение типа эквивалентности.

Что же касается отношения  $\Gamma \cup S$ , то оно не есть отношение типа эквивалентности. Оно соответствует отношению «принадлежать к одному классу», определяемому следующим образом.

**0.4.4.** Два слова  $x$  и  $y$  принадлежат одному  $\Gamma$ - $S$ -классу, если выполнено хотя бы одно из двух следующих условий:

$$\Gamma(x) \cap S(y) \neq \emptyset \text{ или } \Gamma(y) \cap S(x) \neq \emptyset.$$

(Условия, при которых классы образуют разбиение, выясняются в § 34).

Соотношение между классом и разделом таково, что всегда  $\Gamma$ - $S$ -класс целиком входит в раздел.

Соотношение между разделом и типом, вообще говоря, неопределенное, но в реальных языках (а также в моделях, рассматриваемых ниже, см. § 31) раздел совпадает с типом или входит в него.

В русском языке, если считать, что существительные во множественном числе входят в разные семейства, т. е. если учитывать контексты, включающие *одно из которых*, *одна из которых*

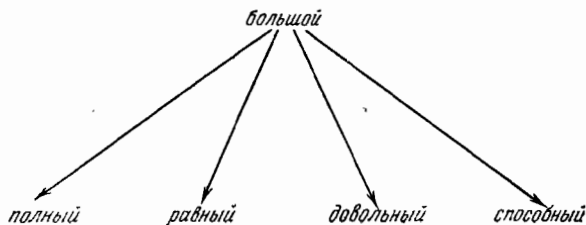
и т. п. (ср. § 22) слова *окно*, *дом*, *книга* входят в разные разделы, но в один тип. Если же рассматривать фрагмент, не содержащий таких контекстов, и считать, что слова *бкна*, *домá*, *книги* входят в одно семейство, то раздел, по-видимому, совпадет с типом.

В сербском языке, где на любом фрагменте не происходит склеивания во множественном числе, раздел всегда уже типа. Если «тип» это в известной мере аналог части речи, то «раздел» и *G-S*-класс суть аналоги объединения слов, вообще говоря более широкого, чем какая-нибудь категория (см. § 39), но более узкого, чем часть речи. В один раздел объединяются группы слов, имеющие хотя бы одну пару эквивалентных форм. Например, если рассмотреть такой подфрагмент только что упомянутого фрагмента русского языка, который содержит лишь существительные в единственном числе, то мы получим два раздела: один, соответствующий существительным женского рода, а другой, соответствующий существительным мужского и среднего родов (*G-S*-цепь, соединяющая слова *мальчик* и *кольцо* на этом фрагменте, имеет вид: *мальчик* G *мальчику* S *кольцу* G *кольцо*). Иначе говоря, для того, чтобы две парадигмы принадлежали одному разделу, достаточно хотя бы одной нейтрализации каких-либо форм этих парадигм.

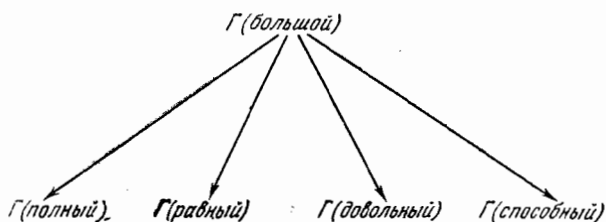
## § 27. ПОНЯТИЯ ГИПЕРТИПА И Т-РЯДА

Покажем теперь, что в ряде случаев понятие типа по объему уже традиционного понятия части речи. Возьмем прилагательное *довольный* и цепочку *довольный завтраками мальчик*. В этой цепочке *довольный* нельзя заменить на *большой* без изменения правильности. Подобное положение наблюдается со всеми прилагательными, имеющими какое-либо управление (например, *равный кому-либо*, *полный чего-либо*, *способный к чему-либо* и т. п.).

Верно однако то, что прилагательное *большой* и любое прилагательное того же типа можно заменить без изменения правильности в любом контексте на *полный*, *равный*, *довольный*, *способный* и т. п.; иначе говоря, имеет место отношение замещения



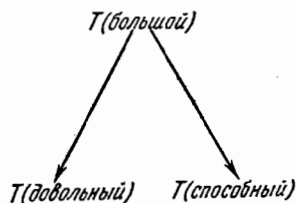
Легко проверить, что подобное отношение имеет место для любой формы соответствующей парадигмы, а именно верно



где парадигмы  $\Gamma$  (полный),  $\Gamma$  (равный) и  $\Gamma$  (довольный) являются  $\Gamma$ -эквивалентными, ибо то, что соответствующие слова *полный*, *равный* и *довольный* управляют разными падежами, не затрагивает соответствующих  $\Gamma$ -образов. В самом деле, возьмем любую пару фраз, например, *довольный жизнью человек* и *полный жизни человек*. Их  $\Gamma$ -образы имеют вид:

$\Gamma$ (довольный)	$\Gamma$ (жизнь)	$\Gamma$ (человек)
$\Gamma$ (полный)	$\Gamma$ (жизнь)	$\Gamma$ (человек)

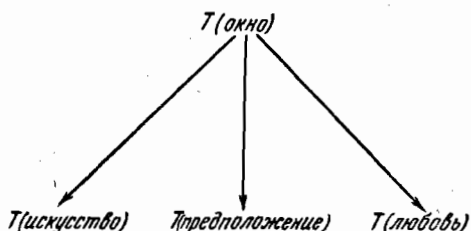
т. е.  $\Gamma$  (довольный) и  $\Gamma$  (полный) всегда встречаются в одинаковых  $\Gamma$ -контекстах. Что касается  $T$  (способный) — то это отдельный тип, поскольку *способный* управляет инфинитивом (а также предлогами *к* и *на*). Заметим, что каждое предложное управление определяет свой тип, аналогичный выделенному типу  $T$  (способный), но мы не будем специально отражать эти типы на схеме. Схема основных  $T$ -замещений для прилагательных имеет вид:



Аналогичное положение с существительными. Поскольку каждое существительное может управлять родительным падежом, то хотя, например, *овладение* и *окно* входят в разные семейства, но  $\Gamma$  (овладение)  $\sim$   $\Gamma$  (окно). Отдельные типы образуют однако

$T$  (искусство) — возможность управлять инфинитивом (мы считаем, что инфинитив и имя действия входят в разные парадигмы, а также  $T$  (предположение) — возможность управлять придаточным предложением с *что*. Кроме того, каждое предложное управление определяет особый тип, например,  $T$  (способность), ср. *способность к*, или  $T$  (объявление), ср. *объявление о*, или  $T$  (любовь),

ср. *любовь к*. На следующей схеме мы отразим лишь один из таких типов



Построим теперь новые аналоги части речи, основанные на идее одностороннего  $T$ -замещения. Пусть все множество  $V$  разбито на типы и на множестве  $T$ -образов фраз определено отношение замещения. Это дает граф в смысле определения 2. 15.

0. 4. 5. Назовем  $\gamma$  и  $\pi$  е р т и п о м слова  $x$  (обозначается  $H(x)$ ) ту часть множества типов, которая включает  $T(x)$  и соответствует максимальному связанному подграфу графа  $T$ -замещений.

Наши примеры показывают, что в русском языке в  $H$  (*окно*) войдут все существительные, а в  $H$  (*большой*) все прилагательные.

Заметим, что для приводимых примеров соответствующие объединения слов можно получить и при помощи другой модели. Вспомним определение 3. 40, которое вводило понятие ряда. Выберем на множестве типов все начальные типы (см. 0. 3. 39).

0. 4. 6.  $T$ - р я д о м назовем множество слов, соответствующее ряду, порожденному некоторым начальным типом.

$T$ -ряды уже не образуют разбиения как гипертипы, но, сколько можно судить по приведенным примерам, для существительных и прилагательных русского языка оба понятия совпадают. Разберем теперь один случай, для которого, по-видимому, понятие  $T$ -ряда имеет некоторые преимущества. С разбираемой точки зрения интересно рассмотреть вопрос об аналоге части речи «наречие».

Посмотрим, на какие семейства распадаются слова *хорошо, плохо, медленно, быстро, лучше, хуже, медленнее, быстрее, там, здесь, весьма, очень, зловеще, мертвецы, по-домашнему*. Допускающие их контексты это такие контексты как:

Контексты  
(Я гуляю,  $\Delta$ )

(я гуляю, чем он)  
(мне, было гулять)  
( $\Delta$ , Москва)  
( $\Delta$ , странная книга)

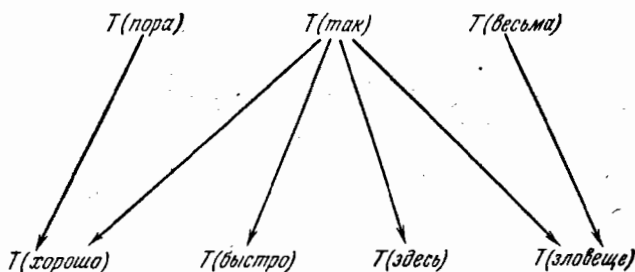
Допускаемые слова

*хорошо, плохо, так, медленно, быстро, там, здесь, зловеще, мертвецы, по-домашнему*  
*лучше, хуже, быстрее, медленнее*  
*хорошо, плохо*  
*там, здесь*  
*весьма, очень, зловеще, мертвецы, по-домашнему.*

Если принять данное здесь (и, по-видимому, достаточно спорное) распределение по контекстам, то получится шесть семейств и шесть типов <sup>4</sup>:

$T_1$ — хорошо, плохо, хуже, лучше	$S_1$ — хорошо, плохо
$T_2$ — медленно, быстро, медленнее	$S_2$ — так, медленно, быстро
$T_3$ — там, здесь	$S_3$ — там, здесь
$T_4$ — зловеще, мертвецки, по-домашнему	$S_4$ — зловеще, мертвецки, по-домашнему
$T_5$ — весьма, очень	$S_5$ — весьма, очень
$T_6$ — так	$S_6$ — хуже, лучше, мед- леннее, быстрее

Переходя к схемам замещения, мы должны заметить, что  $T_1$  замещается еще одним типом, включающим слова *пора*, *можно*, *нужно* и т. п. Схема имеет вид:



Здесь получается всего один гипертип, соответствующий части речи «наречие». Но если мы применим понятие  $T$ -ряда, то получится три разных  $T$ -ряда, соответствующих трем начальным типам. Тот же результат получится, если мы возьмем в качестве аналога части речи гипертип, но предварительно расщепим омонимы второго рода в смысле определения 3.44. В § 15 мы договорились в сущности разделить *хорошо*<sub>1</sub> и *плохо*<sub>1</sub>, употребляемое в таких контекстах, как (*мне, было гулять*) и *хорошо*<sub>2</sub>, *плохо*<sub>2</sub>,

<sup>4</sup> Легко видеть, что типы не совпадают с семействами, если считать, что *хорошо* и *лучше*, *плохо* и *хуже*, *медленно* и *медленнее*, *быстро* и *быстрее* входят в один тип, в то время как *так*, *медленно* и *быстро* образуют одно семейство (последнее спорно). Этот пример чрезвычайно интересен для § 31, где выдвигается требование, чтобы два слова, относящиеся к одному семейству, входили в один тип. Здесь это требование нарушается дважды.



употребляемое в контекстах (*я гуляю*,  $\Lambda$ ) и т. п. Аналогичное разграничение нужно провести и со словами *зловеще*, *мертвецки*, *по-домашнему*, а именно по контекстам (*я гуляю*,  $\Lambda$ ) и ( $\Lambda$ , *странная книга*). Тогда получим гипертипы:

1. *пора*, *надо*, *можно*, *хорошо*<sub>1</sub>, *плохо*<sub>1</sub>
2. *Так*, *медленно*, *быстро*, *там*, *здесь*, *хорошо*<sub>2</sub>, *плохо*<sub>2</sub>, *зловеще*<sub>1</sub>, *мертвецки*<sub>1</sub>, *по-домашнему*<sub>1</sub>.
3. *Весьма*, *очень*, *зловеще*<sub>2</sub>, *мертвецки*<sub>2</sub>, *по-домашнему*<sub>2</sub>.

Что касается выделения двух первых частей речи с соответствующим разграничением омонимов, то такая процедура проводилась рядом русских лингвистов, выделявших специальную часть речи «категория состояния»<sup>5</sup>.

Что касается второго разделения, то оно в какой-то мере противоречит лингвистической интуиции, ибо в славянских языках связи между двумя последними группами очень тесны. Отчасти это связано и с отмеченной выше спорностью выделения допустимых контекстов.

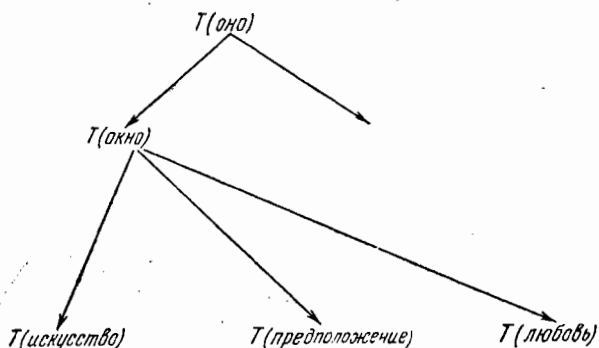
Мы считали, что *я гуляю* очень запрещенная цепочка, следуя здесь мнению Пешковского, писавшего: «Не со всяким глаголом всякое такое наречие можно соединить: нельзя сказать *очень читает*, *очень гуляет*, *очень спит*, *весьма лежит* и т. д. (говорится иногда в шутку) . . .»<sup>6</sup> С другой стороны, для того, чтобы цепочка *Я гуляю* очень была фразой, достаточно, чтобы она не была запрещена (ибо *очень* заменяемо, например, с *хорошо* в других контекстах), т. е. в интерпретации была потенциально возможной, а последнее, по-видимому, не исключено. Если это так, то останутся всего два *T*-ряда, из которых один соответствует собственно наречиям, а другой так называемой «категории состояния».

Мы видим, таким образом, что, приняв понятие гипертипа, мы получаем, с одной стороны, интересные и оправдываемые интуитивно выводы, а, с другой стороны, должны считаться с возможностью результатов, непредвиденных заранее. Это особенно проявляется, когда мы рассматриваем вопрос о том, к каким гипертипам относятся местоимения. Очевидно, что в любом контексте местоимение-существительное можно заменить на существитель-

<sup>5</sup> Наиболее удачно этот вопрос изложен в статье Л. В. Щербы «О частях речи в русском языке» («Русская речь», новая серия, II. Л., 1928; [перепечатано в кн.] Л. В. Щерб а. Избранные работы по русскому языку. М., 1957).

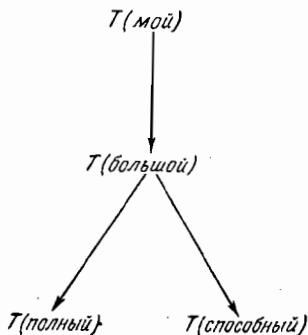
<sup>6</sup> А. М. Пешковск и й. Русский синтаксис в научном освещении. Изд. 6. М., 1938, стр. 118.

ное. Это означает, что на самом деле гипертип, соответствующий существительному, имеет следующую структуру:



Это должно интерпретироваться так, что наиболее характерным существительным с синтаксической точки зрения, представленной вообще в нашей работе, является местоимение-существительное. Такой вывод интересно сопоставить с мыслями Пешковского о том, что в словах *кто* и *что*, относящихся, как мы видели к *T(она)*, «выступают не сами предметы, а как бы одна «чистая предметность», «чистая существительность», это идеальные существительные»<sup>7</sup>.

Аналогично любое местоимение-прилагательное, например, *мой*, *этот*, *какой* можно заменить на какое-нибудь прилагательное, например, *богатый*. Обратная замена невозможна в контекстах, содержащих *очень*, *весьма*, а также в контекстах уже содержащих местоимения-прилагательные, например, в *мой прекрасный город* нельзя заменить *прекрасный* на местоимение-прилагательное. Вновь приходится дополнить схему, приведенную выше:



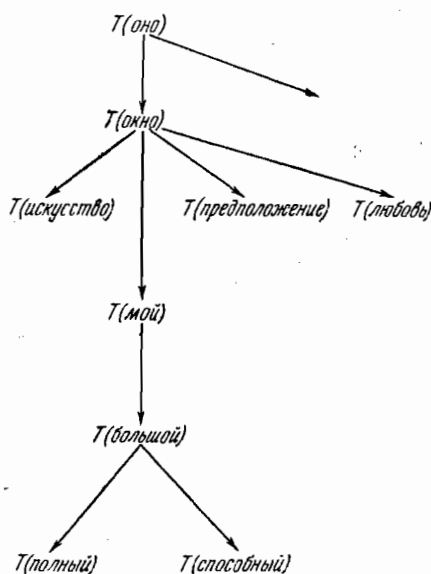
<sup>7</sup> А. М. Пешковский. Русский синтаксис в научном освещении, стр. 94.

И здесь можно, повторив сказанное выше, сослаться на Пешковского, который считал слово *какой* (относящееся по нашей процедуре к *T* (*мой, этот*) «идеальным прилагательным»<sup>8</sup>).

Мы видим, что в нашей классификации не получается специальной части речи «местоимение».

Особенность предложенной классификации состоит в ее зависимости от выбранного множества реальных и запрещенных цепочек (а тем самым и от множества фраз).

Если допустить в число реальных фраз и такие эллиптические фразы, как: *я вижу красивую, я вышел в большую, красивая сказала* и т. п., то тогда окажется, что в произвольном контексте можно заменить существительное на прилагательное, но не наоборот и схемы на стр. 105 объединяются в следующей схеме:



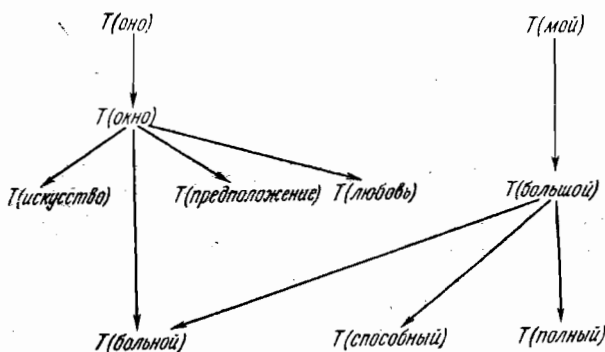
Таким образом, мы получаем единый гиперттип имен.

Мы будем однако считать, что одновременно рассматриваются два задания реальных фраз (при одном эллиптические исключаются, при другом включаются).

Полностью исключить эллиптические фразы рассматриваемого вида нельзя из-за принципиальной возможности субстантивации любого прилагательного. Осуществление этой возможности зависит скорее не от внутренней грамматической природы явлений, а от внеязыковой действительности (которая и закреплена в реальности тех или иных фраз). Разумеется, если бы был всего один

<sup>8</sup> А. М. Пешковский. Русский синтаксис в научном освещении, стр. 103.

промежуточный тип прилагательно-существительных, например, таких как *больной*, *знакомый* и т. п., и допускались бы лишь фразы вида *я вижу знакомую*, *я вошел в столовую*, *больная сказала*, в то время как цепочки *я вижу красивую*, *я вошел в большую* не были бы фразами, то можно было бы предложить более естественную схему:



Тогда *T(больной)* относилось бы к двум *T*-рядам, но к одному гипертипу. Это соответствовало бы двойной (с одной стороны, контекстной и, с другой стороны, словарной) характеристике таких слов, т. е. случаю категориальной омонимии<sup>9</sup>.

К сожалению, подобное решение слишком упрощает языковую действительность. Как указывает Яодловский<sup>10</sup>, можно выделить по крайней мере 5 ступеней перехода от прилагательного к существительному:

1. слова, не подвергающиеся процессу субстантивации — польск. *drewniany* (русск. *деревянный*), *szorsty*, *dokuczliwy*;
2. начало процесса субстантивации — польск. *chory* (ср. русск. *больной*), *pijany* (ср. русск. *пьяный*), *ubogi*, *szalony*;
3. состояние равновесия — польск. *znajomy* (ср. русск. *знакомый*), *łysy* (ср. русск. *лытый*), *gluchoniemy* (ср. русск. *глухонемой*), *zmarły* (ср. русск. *мертвый*);
4. перевес субстантивации — *podroźny*, *personalny*;
5. полный переход в существительные — *leśniczy* (ср. русск. *лесничий*), *księgowy*.

Слова второй группы в отрыве от контекста воспринимаются как прилагательные, слова третьей группы в отрыве от контекста воспринимаются как промежуточные образования, а слова четвертой группы скорее воспринимаются как существительные.

<sup>9</sup> См.: S. Jodłowski. O kontekstowym i słownikowym kwalifikowaniu części mowy. «Biuletyn polskiego towarzystwa językoznawczego». Zeszyt XX, Fasc. XX, 1961, стр. 60—62.

<sup>10</sup> Jodłowski. Указ. соч., стр. 62; ср. также: O. Leška. K substantivaci adjektiv. «Bulletin ústavu ruského jazyka a literatury, 1965, IX, стр. 13—17, где выделяются аналогичные группы.

Очевидно, что то или иное восприятие зависит от числа реальных контекстов, в которых соответствующее слово употребляется. Однако в моделях, не учитывающих статистических соображений, мы волей-неволей должны ограничиться фиксированием потенциальных возможностей.

## § 28. ИСПОЛЬЗОВАНИЕ ИДЕИ СИНТАКСИЧЕСКОЙ ОБУСЛОВЛЕННОСТИ ПРИ МОДЕЛИРОВАНИИ ЧАСТИ РЕЧИ

Мы видели, что для некоторых пар слов, относящихся к одной части речи, а именно таких, сочетательные способности которых резко различаются (например, возможность или невозможность сочетаться с придаточным предложением) понятие типа становится слишком узким, что привело нас к понятиям гипертита и  $T$ -ряда.

В других случаях понятие типа оказывается слишком широким и в один тип могут попасть в русском языке слова, относящиеся к разным частям речи. Связано это со следующим обстоятельством. Переходя к  $G$ -образам, мы отвлекаемся от всех синтаксических зависимостей между словами, выраженных формой слов (т. е. выбором определенного члена парадигмы). Есть, правда, и другой способ выражения синтаксических зависимостей, а именно порядок слов, но как раз в славянских языках в связи с развитой системой флективных форм порядок слов свободный (почти не знающий ограничений, если брать изолированные предложения).

Более того, известно, что оба средства выражения синтаксической зависимости дополняют друг друга в том смысле, что фиксированный порядок слов как правило встречается в языках с слабо развитой системой формообразования (есть и исключения, например, тюркские языки) и наоборот.

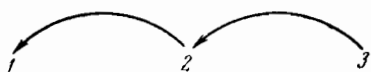
Дело объясняется, однако, тем, что, как часто бывает в лингвистической теории моделей, не была выяснена целиком интуиция, давшая повод для формирования соответствующего понятия. А именно взаимозамещаемость практически проверялась лишь для тех  $G$ -образов, которые имели одинаковые синтаксические связи между парадигмами. Иначе говоря, отвлекаясь от специфической формы слова во фразе, исследователь не отвлекался от связей этого слова с другими словами, но в модели это обстоятельство не было оговорено. Мы попробуем теперь отразить в модели и это понятие.

0.4.7. Будем говорить, что фраза  $f = x_1 \dots x_n$  имеет ту же синтаксическую структуру, что и фраза  $g$ , если множества вхождений в эти фразы изоморфны в следующем смысле: для любых  $i$  и  $j$  слово  $x_i$  синтаксически подчиняет  $x_j$ , если и только если  $y_i$  синтаксически подчиняет  $y_j$ .

Примеры фраз, имеющих одинаковую синтаксическую структуру:

*Маленькая<sub>1</sub> девочка<sub>2</sub> лежит<sub>3</sub>*

*Его<sub>1</sub> брат<sub>2</sub> спит<sub>3</sub>*



0.4.8. Мы будем говорить, что парадигма  $\Gamma(x)$  замещает с учетом синтаксического подчинения  $D$ , или короче  $D$ -замещает парадигму  $\Gamma(y)$ , и записывать это через  $\Gamma(x)/\Gamma(y)$ , если для любого контекста  $(z_1 \dots z_k, z_{k+1} \dots z_n)$ , полученного из некоторой фразы  $f$  путем вычеркивания слова  $x' \in \Gamma(x)$ , найдется контекст  $(z'_1 \dots z'_k, z'_{k+1} \dots z'_n)$  такой, что а) он получен из фразы  $f'$ , имеющей ту же синтаксическую структуру, что и  $f$ , вычеркиванием слова  $y' \in \Gamma(y)$ ; б) для любого  $i$  верно  $z'_i \in \Gamma(z_i)$ .

0.4.9. Мы будем говорить, что  $x$  и  $y$  принадлежат к одному синтаксически обусловленному типу  $\tau$ , если имеет место  $\Gamma(x)/\Gamma(y)$  и  $\Gamma(y)/\Gamma(x)$ .

0.4.10. Пусть  $\tau_1$  и  $\tau_2$  два синтаксически обусловленных типа. Мы будем говорить, что  $\tau_1$  квалифицированно замещает  $\tau_2$ , если существует  $\Gamma_1 \in \tau_1$  и  $\Gamma_2 \in \tau_2$  такие, что  $\Gamma_1/\Gamma_2$ .

Примечание. Методом, примененным при доказательстве теоремы 3.3, можно показать, что такое определение вполне корректно, ибо в этих условиях любая парадигма из  $\tau_1$   $D$ -замещает любую парадигму из  $\tau_2$ .

0.4.11. Возьмем граф отношения  $D$ -замещения синтаксически обусловленных типов. Максимальный связный подграф этого графа назовем синтаксически обусловленным гипертипом.

Можно было бы аналогичным способом ввести понятие, близкое  $T$ -ряду, но мы этого делать не будем.

Нам хочется указать на следующие обстоятельства. Ясно, что введение условия тождественности синтаксической структуры накладывает более сильные требования на принадлежность двух слов к одной части речи, но в остальном оставляет структуру понятий неизменной.

Для интерпретации модели на фактах славянских языков введенные более сложные понятия удобней, и в сущности при анализе примеров мы неявно пользовались идеей синтаксической обусловленности. Так, в рассуждениях на стр. 101 мы неявно предполагали, что  $\Gamma$  (полный) не замещает  $\Gamma$  (большой), ср. однако неограниченную возможность таких сопоставлений, как  $\Gamma$  (полный)  $\Gamma$  (жизнь)  $\Gamma$  (студент) и  $\Gamma$  (большой)  $\Gamma$  (жизнь)  $\Gamma$  (студент), где в первом случае выписан  $\Gamma$ -образ фразы

полный жизни студент, а во втором Г-образ фразы большая жизнь студента. Теперь такие сопоставления невозможны, ибо при этом меняется направление синтаксического подчинения.

Может возникнуть вопрос, почему же мы с самого начала не вводим синтаксической обусловленности типа и гипертипа? Нам представляется наиболее важным не самый аналог части речи, а процесс его получения, процесс постепенного усложнения модели с целью все более адекватного отражения языковой действительности.

Именно моделирование части речи оказывается с этой точки зрения особенно показательным, вводится сначала тип, затем модель усложняется рассмотрением одностороннего замещения и, наконец, происходит введение идей синтаксической обусловленности части речи. Последнее важно с методологической точки зрения и как иллюстрация выдвинутого в § 3 тезиса о необходимости связи идей моделирования с лингвистической традицией. Первые аналоги «части речи» совершенно не учитывали работ, весьма важных для общего развития советского языкознания и приведших к идее, что части речи понятие вторичное по отношению к идее синтаксических зависимостей<sup>11</sup>. Эта идея явным образом использована в модели этого параграфа.

Использование идеи синтаксической обусловленности, по-видимому, наиболее плодотворно при обсуждении части речи «глагол» (для которой как раз характерно, что она вступает в наибольшее число синтаксических связей).

## § 29. ЧАСТЬ РЕЧИ «ГЛАГОЛ». НЕКОТОРЫЕ ОБЩИЕ ЗАМЕЧАНИЯ О МОДЕЛИРОВАНИИ ЧАСТИ РЕЧИ

По связям глагола с субъектом, а также с инфинитивом и придаточными предложениями и объектами, можно выделить, например, следующие группы глаголов<sup>12</sup> (см. таблицу на стр. 111).

Каждая такая группа соответствует синтаксически обусловленному типу, так как различия в управлении конкретными падежами (например, для *мне нездоровится* и *я кашляю*) несущественно. Слова двух разных групп соответствуют двум разным типам, например, *рассказывает* и *колет*, ибо для Г-образа фразы *он рассказывает ему, что болен* нельзя подобрать такой Г-образ, который отличался бы только заменой Г (*рассказывает*) на Г (*колет*).

Для того чтобы построить схему синтаксически обусловленных τ-замещений, мы должны условиться, какие опущения объек-

<sup>11</sup> См., например: И. И. Мещанинов. Части речи и члены предложения. М., 1945.

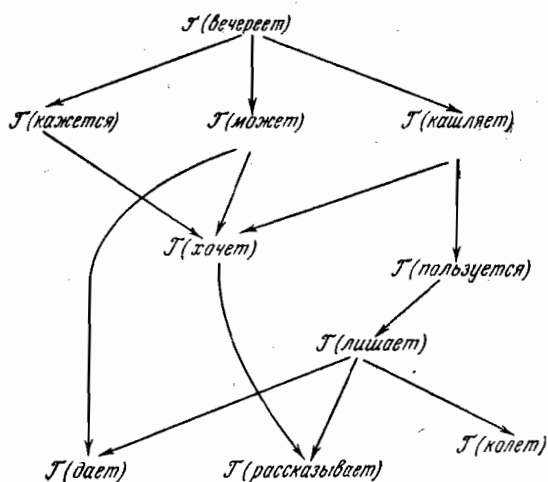
<sup>12</sup> Ср. классификацию глаголов славянских языков: Р. Мразек. Синтаксическая дистрибуция глаголов и их классы. — ВЯ, 1964, № 3, стр. 50—62.

Пример	Субъект	Число объектов существительных	Инфинитив	Придаточное
1. <i>вечерет, брезжит</i>	—	0	—	—
2. <i>нездоровится, кашляет, спит</i>	+	0	—	—
3. <i>кажется</i>	+	0	—	+
4. <i>хочется, хочет</i>	+	1	+	+
5. <i>может</i>	+	0	+	—
6. <i>пользуется, использует</i>	+	1	—	—
7. <i>дает</i>	+	2	+	—
8. <i>лишает</i>	+	2	—	—
9. <i>рассказывает, считает</i>	+	2	—	+
10. <i>колет, несет</i>	+	3	—	—
11. <i>пишет</i>	+	3	—	+

тов в этой схеме считаются еще допустимыми. Если считать, что любое опущение объектов возможно, т. е. например, реальны все фразы, приводимые ниже

*он рассказывает ему сказку*  
*он рассказывает сказку*  
*рассказывает сказку*  
*рассказывает*

то схема синтаксически обусловленных  $\tau$ -замещений будет иметь вид





Может быть, следует допускать в качестве эллиптических лишь фразы с полнозначными глаголами. Тогда такие глаголы, как *мочь, быть, казаться*, попадут в разные синтаксически обусловленные гипертипы (аналогично с другими служебными словами) и не будут связаны с полнозначными глаголами. Все это вопросы, которые теория моделей решить не может, она проясняет лишь связь между той или иной классификацией частей речи и заданием множества реальных фраз.

Можно, по-видимому, предположить, что при одинаковых принципах задания реальных и запрещенных цепочек во всех славянских языках выделяются одинаковые (синтаксически обусловленные) гипертипы и типы. Разница между отдельными славянскими языками начинает проявляться на уровне разделов (для определенных фрагментов) и уже достаточно резко выявляется на уровне категорий рода (ср. главу VI). Поскольку грамматический род, как он моделируется в главе VI, целиком включается в раздел, раздел целиком включается в тип, а тип в гипертип, то вполне оправдано то обстоятельство, что при более грубом объединении слов, различия между языками стираются, в то время как обращение к более тонкой структуре наоборот проясняет соответствующие различия.

**П р и м е ч а н и е.** Интересно, что наша классификация по частям речи действительна только для полнозначных слов, а для неполнозначных неприемлема, так как для последних почти каждое слово выделилось бы в отдельный гипертип. Таким образом, первоначально должно быть произведено деление на служебные и полнозначные слова (о возможности формального выделения служебных слов см. § 30) и лишь потом в пределах неслужебных слов должно производиться разбиение на части речи. В сущности именно такой подход применяется в грамматиках русского и ряда других славянских языков, где разделяются части речи и частицы<sup>13</sup>.

---

<sup>13</sup> См. об этом, например, в кн.: «Основы построения описательной грамматики современного русского литературного языка». М., 1966, стр. 93—96.

ТЕОРЕТИКО-МНОЖЕСТВЕННАЯ  
ТИПОЛОГИЯ ЯЗЫКОВ§ 30. ПОНЯТИЯ АМОРФНОГО  
И ПРЕДЕЛЬНО ГРАММАТИКАЛИЗОВАННОГО ЯЗЫКА

Теперь мы приступаем к выполнению нашей основной задачи, а именно выяснению типологической характеристики славянских языков, т. е. характеристики в терминах теории моделей. Для этого мы попытаемся ввести прежде всего некоторые аналоги понятий, употребляющихся в традиционной типологии. Мы начнем с понятия «аморфного языка», причем сохраним тот же термин, не боясь смешения понятий, ибо, во-первых, термин «аморфный» почти повсеместно оставлен в типологии, а, во-вторых, он очень удобен для следующих определений.

**0. 5. 1.** Язык называется аморфным, если для любых слов  $x$  и  $y$  из того, что  $y \in G(x)$  следует, что  $y = x$ .

Иначе говоря, в аморфном языке каждая парадигма состоит ровно из одного слова, т. е. в таком языке нет форм словоизменения, нет морфологии.

В аморфном языке вся система понятий модели упрощается, а именно имеет место: для любого  $x$  верно  $S(x) = Ra(x) = T(x)$ .

Наряду с понятием аморфного языка определим понятие аморфного слова (идея локализации основных свойств языка принадлежит С. Маркусу)<sup>1</sup>.

**0. 5. 2.** Слово  $x$  называется аморфным, если его парадигма состоит ровно из одного слова.

Очевидно, что язык аморфен тогда и только тогда, когда каждое его слово аморфно.

Исследуя в первую очередь славянские языки, мы заинтересованы не столько в понятии аморфного языка, сколько в понятии аморфного слова. Дело в том, что в славянских языках нельзя выделить даже аморфного подязыка (фразы типа *из метро да в кино* — на основе реальной *из огня да в полымя* или часть реальной фразы *уж замуж невтерпех* явные раритеты). Приведенные примеры интересны лишь тем, что показывают два вида аморфных слов: а) такие слова, как *метро*, которые являясь аморфными, непосредственно связаны со словами, не являющимися аморфными,

<sup>1</sup> S. Marcus. *Lingvistică matematică*. București, 1963.

причем эта связь вскрывает омонимичный характер соответствующих слов и б) аморфные слова, не обладающие таким свойством.

В связи с этим введем следующие понятия.

0. 5. 3. Аморфное слово  $x$  называется собственно аморфным, если не существует двух разных слов  $y_1$  и  $y_2$ , таких, что  $y_1 \in G(y_2)$ ,  $y_1 \Rightarrow x$  и  $y_2 \Rightarrow x$ .

0. 5. 4. Всякое аморфное слово, которое не является собственно аморфным, называется не собственно-аморфным.

Несобственно аморфное слово может получить по аналогии формы словоизменения (ср. *пальто, пальта, пальту, пальте* и т. п.)

Формулировка понятия «аморфности» в терминах модели позволяет поставить следующий вопрос: какими лингвистическими свойствами будет обладать язык, если в определениях 5.1 и 5.2 заменить понятие парадигмы понятием семейства (само определение без постановки вопроса о лингвистических свойствах предложено С. Маркусом <sup>2)</sup>?

Рассмотрим поэтому следующие определения:

0. 5. 5. Язык называется предельно грамматикализованным, если в нем для любых  $x$  и  $y$  из  $y \in S(x)$  следует  $y = x$ .

0. 5. 6. Слово называется предельно грамматикализованным, если его семейство состоит ровно из одного слова.

Термин «предельная грамматикализованность» выбран нами из следующих соображений. Если слово обозначает предметы и явления внешнего мира, то число слов в его семействе как правило очень велико и никогда не задается списком (таковы существительные, прилагательные, полнозначные глаголы, некоторые типы наречий). Если же слово выполняет грамматическую роль (частицы, предлоги, союзы) или даже передает очень общие отношения (наречий степени, модальные и связочные глаголы), то в семействе соответствующего слова очень мало слов. Поэтому такие слова обычно все без исключения приводятся списком в грамматиках. В пределе, который мы и отразили в определении 5. 6, семейство и состоит из одного слова.

Ясно, конечно, что предельно грамматикализованные языки не могут существовать в действительности: это были бы языки без лексики, ибо язык не мог бы функционировать, если для названия каждого нового предмета или явления действительности нужно было бы новое правило грамматики (напомним, что согласно основному тезису теории моделей семейство совпадает с категориальным пучком).

Зато понятие предельно грамматикализованного слова лингвистически вполне прозрачно. Оно может быть использовано для

---

<sup>2</sup> S. Marcus. *Lingvistică matematică*, стр. 184—185.

классификации служебных слов и для определения аналога этого понятия.

0. 5. 7. Называем слово *служебным*, если оно является собственно-аморфным или предельно-грамматикализованным.

Таким образом, служебными словами мы считаем такие собственно-аморфные слова, как частицы и союзы и такие собственно-аморфные и предельно грамматикализованные слова, как предлоги, а также такие предельно грамматикализованные слова, как *я* и *ты*. Правда, в разряд служебных попадут и некоторые неизменяемые наречия (например, *замуж*). Это показывает, что классификация, используемая в этой модели, в известной степени несовершенна, хотя она и отражает некоторые существенные черты описываемых явлений. С другой стороны, можно усилить требование к «служебности», введя следующее определение:

0. 5. 8. Слово называется *предельно-служебным*, если оно является одновременно собственно-аморфным и предельно-грамматикализованным, т. е. из  $y \in S(x) \cup \Gamma(x)$  следует  $y = x$ .

С точки зрения этого определения предельно служебными в славянских языках являются, по-видимому, некоторые предлоги и глагольные частицы типа *да* в болгарском, играющие аналогичную роль. Если бы в качестве исходных элементов брались не слова, а морфы, то предельно-служебными были бы окончания и предлоги, что показывает близость одних к другим.

При определении служебного слова мы нигде не пользовались критерием полнзначности, введенном в § 14.

Нам было важно, что используя лишь понятия семейства и парадигмы можно подойти (хотя и с известными ограничениями, о которых мы говорили), к классификации, обычно использующей критерий полнзначности.

### § 31. ПОНЯТИЕ S-T-ПРАВИЛЬНОГО ЯЗЫКА

Мы привыкли, что в русском языке и других славянских два слова, относящиеся к одному классу взаимозаменяемых словоформ как правило относятся к одной части речи. Более того согласно основному тезису теории моделей семейство должно совпадать с категориальным пучком. Естественно поэтому ожидать, что два слова, входящие в те же самые категории, будут относиться к той же части речи.

В общем случае это не так. Как заметил В. А. Успенский<sup>3</sup>, можно построить пример языка, в котором два слова, относящиеся к одному семейству, входят в разные типы.

В связи с этим было предложено следующее определение.

<sup>3</sup> В. А. Успенский. К определению части речи в теоретико-множественной системе языка. «Бюллетень объединения по проблемам машинного перевода». М., 1957, № 5.

0.5.9. Язык называется *S-T*-правильным, если любые два слова, принадлежащие к одному семейству, принадлежат и к одному типу.

Примечание. Термин «правильный язык», введенный В. А. Успенским, неудачен, ибо ничего не говорит о лингвистических свойствах такого языка. Поэтому параллельно мы предлагаем другой термин: «язык с категориально однозначными частями речи». Правда, этот термин достаточно громоздок.

Имеет место следующая теорема, показывающая отношения между типом и разделом, уже обсуждавшиеся нами на эмпирическом уровне.

Т. 5. 1. В *S-T*-правильном языке для всякого слова  $x$  верно:

а)  $Ra(x) \subseteq T(x)$ , т. е. раздел содержится внутри типа;

б)  $T(x) = Ra^*(x)$ , т. е. разбиение на типы совпадает с производным от разбиения на разделы.

Раньше считалось, что свойством правильности характеризуются любые естественные языки (отсюда и название «правильный язык»).

Сейчас накопился уже целый ряд фактов, свидетельствующих о том, что принцип правильности не является языковой универсалией.

Наиболее показателен пример, приведенный в § 27, где выясняется, что слова *медленнее*, *быстрее*, *лучше*, *хуже* относятся к одному семейству, но к разным типам: Это же верно относительно слов *так* и *медленно*. Правда, можно было бы сказать, что наречия это вообще класс слов, характеризующихся всякими неправильностями, и поэтому мы имеем здесь дело с периферийным явлением. Однако из других языков можно привести примеры неправильности, проявляющейся и в более центральных типах слов.

Таким примером нарушения правильности (в смысле приведенного определения) являются некоторые факты болгарского языка. Известно, что в болгарском языке степени сравнения могут быть образованы не только от прилагательных, но и от некоторых существительных, ср.:

<i>добър</i>	— <i>пб-добър</i>	‘хороший — более хороший’
<i>юнак</i>	— <i>пб-юнак</i>	‘герой — более герой’
<i>човек</i>	— <i>пб-човек</i>	‘человек — более человек’
<i>майстор</i>	— <i>пб-майстор</i>	‘мастер — более мастер’

Если, например, считать, что *пб-юнак* есть форма слова *юнак* (на что, по-видимому, указывает ударение на *пб*), то поскольку *пб-юнак* употребляется на определенном фрагменте так же, как, например, слово *подобър*, эквивалентными оказываются словоформы, относящиеся к разным частям речи.

Типологически сходные явления представляют сравнительные формы существительных в финно-угорских языках, ср. эст. See

mees on velgi ahvim 'этот человек еще более обезьяна' (ahv — 'обезьяна'; im — суффикс сравнительной степени).

Приведенные примеры показывают, что целесообразно локализовать определение правильности, отнеся его не к языку вообще, а к отдельным словам.

0. 5. 10. Слово  $x$  называется  $S$ - $T$ -правильным, если любое слово  $y$ , входящее в то же семейство, относится к тому же типу.

Очевидно, что имеет место следующее утверждение.

Т. 5. 2. Язык  $S$ - $T$ -правилен тогда и только тогда, когда каждое слово  $x$ , входящее в его словарь, является  $S$ - $T$ -правильным (это утверждение непосредственно следует из определений).

Применительно к приведенным болгарским примерам можно сказать, что слова *юнак*, *майстор*, *човек*, а также все качественные прилагательные неправильны, но, по-видимому, остальные слова этого языка являются  $S$ - $T$ -правильными в смысле нашего определения.

Понятие правильности сформулировано применительно к двустороннему замещению. Маркус в цитированной работе предложил ввести более сильное условие, учитывающее и случаи одностороннего замещения.

0. 5. 11. Язык называется вполне правильным, если в нем для любых слов  $x$  и  $y$  из  $x \Rightarrow y$  следует  $x \in T(y)$ . Слово  $x$  называется вполне правильным, если для любого слова  $y$  такого, что  $y \Rightarrow x$ , верно  $x \in T(y)$ .

## § 32. ПРОСТОЙ ЯЗЫК И НОРМАЛИЗОВАННЫЙ ЯЗЫК

0. 5. 12. Слово  $x$  называется простым, если для любых  $x$  и  $y$  из  $y \in T(x) \cap S(x)$  следует, что  $y = x$ ; язык называется простым, если каждое его слово просто.

Понятие простоты, введенное О. С. Кулагиной<sup>4</sup>, покрывает два ряда совершенно разнородных с лингвистической точки зрения явлений. Прежде всего, сюда относится наличие вариантных форм типа *профессора*—*профессоры* в русском языке. Таких форм в любом славянском языке очень много (все случаи колебания ударений в русском языке, параллельные формы в современном чешском языке и т. п.). Они чрезвычайно важны для развития языка, с ними на каждом шагу встречаются диалектология и стилистика. Сравнивать языки по наличию таких форм было бы неправильно, ибо число их непосредственно зависит от уровня нормализации, стандартизации языка. Поэтому мы объединим подобные формы, используя отношение свободного варьирования. Заменяем в определении 2.38 «звук» на «слово», «слово» на «фразу», а «дезигнат» на «смысл». Можно ввести следующее определение.

<sup>4</sup> О. С. Кулагина. Об одном способе определения грамматических понятий на базе теории множеств. «Проблемы кибернетики», вып. 1. М., 1958.

0. 5. 13. Два слова  $x$  и  $y$  называются грамматическими вариантами друг друга, если они входят в одну парадигму и находятся в отношении свободного варьирования. Язык, в котором нет грамматических вариантов (т. е. варианты считаются разными манифестациями одного знака-типа в смысле § 1) назовем нормализованным.

Мы ограничимся в дальнейшем только нормализованным языком. Здесь мы можем более четко зафиксировать другое проявление непростоты, которое для структурной типологии гораздо более интересно.

В качестве примера рассмотрим определенные и неопределенные имена в болгарском языке. Здесь параллельно употребляются определенные и неопределенные формы в одном и том же контексте (пример Андрейчина<sup>5</sup>):

*Днес не съм чел вестник*

*Днес не съм чел вестника*

Вообще говоря, если не считать некоторых фразеологических сочетаний, определенные формы в любом контексте можно заметить на неопределенные и обратно без изменения грамматической правильности, причем они не являются вариантами друг друга, так как смысл фразы при замене одного слова на другое меняется. А это означает, что соответствующие слова не являются простыми. Это, как мы увидим в следующей главе, сильно ограничивает возможности моделирования агглютинативно выражаемых категорий.

Обобщим теперь понятие простоты на случай одностороннего замещения.

0. 5. 14. Слово  $x$  называется вполне простым, если для любого  $y$  ( $y \neq x$ ) из  $x \Rightarrow y$  или  $y \Rightarrow x$  следует, что  $x$  не входит в  $\Gamma(y)$ . Язык вполне прост, если любое слово его словаря вполне просто.

Мы видели, что большинство слов русского языка суть простые слова. По-видимому, усиление требования простоты не очень меняет положение в русском языке. В пределах парадигмы почти не встречается такого положения, когда для двух слов парадигмы  $x$  и  $y$  имеет место  $y \Rightarrow x$  или  $x \Rightarrow y$ .

Положение резко меняется если мы транзитивируем отношение замещения в определении 5. 14.

0. 5. 15. Слово  $x$  называется идеально простым, если для любого  $y$  ( $y \neq x$ ), если  $y \in D(x)$ , то  $y$  не входит в  $(\Gamma x)$ . Язык идеально прост, если каждое слово его словаря идеально просто.

Идеально простых слов в славянских языках, по-видимому, не больше, чем не идеально простых. В самом деле такие слова как *зиме* или *зимы* не идеально простые, ибо имеют место отно-

<sup>5</sup> Л. Андрейчин. Грамматика болгарского языка. Перев. с болгарского. М., 1949, стр. 97.

шения замещения *зиме*  $\Rightarrow$  *ночи*  $\Leftarrow$  *зимы*, а стало быть *зиме* и *зимы* входят в один  $D$ -класс. Поскольку внутрипарадигматическая омонимия (омонимия первого рода) очень распространена в славянских языках, то их важным типологическим свойством является отсутствие идеальной простоты. Внутрипарадигматической омонимии нет в агглютинативных языках, но там, как мы видели, простота нарушается по другой причине. Что касается аморфных языков, то они, разумеется, идеально простые. Но это, так сказать, языки с нулевой парадигматикой. Для того чтобы отметить это свойство, введем еще одно определение.

0.5.16. Языком с ненулевой парадигматикой является язык, в котором для любых  $x$  и  $y$  таких, что  $x \Rightarrow y$ , найдутся слова  $x' \in \Gamma(x)$  и  $y' \in \Gamma(y)$  такие, что  $x' \in S(y')$ ; словом с ненулевой парадигматикой называется любое слово  $x$  такое, что если существует  $y \Rightarrow x$ , то найдутся  $x' \in \Gamma(x)$  и  $y' \in \Gamma(y)$  такие, что  $x' \in S(y')$ .

Таковыми в русском языке являются так называемые неизменяемые существительные: *метро*, *пальто*, *бордо*, *Ницца* и т. п. В самом деле, мы видели, что имеет место *окно*  $\Rightarrow$  *метро*, но в парадигме слова *окно* нет слова, эквивалентного слову *метро*.

Интересно, что в чешском языке слов с ненулевой парадигматикой значительно меньше, чем в русском (такими являются лишь слова типа *Bordeaux*, *Sanssouci* и т. п.).

Наконец, типологической особенностью сербского и словенского языков является тот факт, что в них, по-видимому, вообще нет слов с нулевой парадигматикой, и, в частности, все иностранные слова имеют столько же форм, что и исконные слова языка.

### § 33. ПОНЯТИЕ ПОЛНОГО ЯЗЫКА

0.5.17. Язык называется п о л н ы м, если в нем для любых  $x$  и  $y$  из  $x \Rightarrow y$  следует, что любой контекст, допускающий  $y$ , допускает и  $\Gamma(x)$ .

Вообще говоря, славянские языки не являются полными. Например, имеем отношение *окно*  $\Rightarrow$  *овладение*, но контекст (*борьба за, знаниями*) не допускает парадигмы  $\Gamma(\text{окно})$ .

Тем не менее понятие полного языка имеет определенное значение.

В самом деле, если ограничиться зависимостями вроде *окно*  $\Rightarrow$  *метро* или *юноши*  $\Rightarrow$  *мальчика*, то очевидно, что они полноты не нарушают.

### § 34. ПОНЯТИЕ ОДНОРОДНОГО ЯЗЫКА

В модели О. С. Кулагиной введено одно чрезвычайно важное понятие, которое мы будем называть свойством формальной однородности языка (или короче однородности).



0. 5. 18. Язык называется формально однородным, если в нем для любых  $x$  и  $y$  из  $\Gamma(x) \cap S(y) \neq \emptyset$  следует

$$\Gamma(y) \cap S(x) \neq \emptyset.$$

В модели формально однородного языка доведены до завершения те тенденции выравнивания по аналогии, которые действуют в каждом языке и о которых так много писалось в лингвистике. Очень близки к этой модели агглютинативные языки (в особенности, система имени тюркских языков), а также такие искусственные языки как эсперанто. Осуществление выдвинутых Кулагиной абстрактных требований в эсперанто показывает, что интуитивно они всегда ощущались.

Понятие формальной однородности обладает и важными теоретическими свойствами, а именно с его помощью можно доказать теоремы:

Т. 5. 3. В формально однородном языке раздел совпадает с  $\Gamma$ - $S$ -классом.

Т. 5. 4. Всякий формально однородный язык есть правильный язык.

Значение модели формально однородного и простого языка можно проиллюстрировать еще одним примером. Служебные слова, как мы уже говорили, аморфны в смысле определения 5. 2. Имеет место следующая теорема:

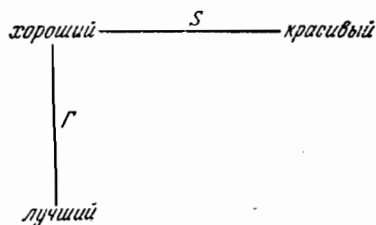
Т. 5. 5. В простом и формально однородном языке каждое слово,  $E$ -эквивалентное аморфному слову, само является аморфным.

Отклонения от формальной однородности наблюдаются довольно часто и могут быть использованы для типологического сравнения языков. Здесь также полезно локализовать понятие формальной однородности:

0. 5. 19. Слово  $x$  называется формально однородным, если для любых слов  $y$  и  $z$  таких, что  $x \in \Gamma(y) \cap S(z)$  найдется слово  $w$  такое, что  $w \in \Gamma(z) \cap S(y)$ .

Очевидно, что язык формально однороден тогда и только тогда, когда каждое его слово формально однородно.

Примером формально неоднородных слов русского языка могут служить слова *хороший*, *плохой*, *высокий*, *низкий*, *маленький*, *большой*. Если считать, например, что *лучший* и *хороший* суть формы одного слова (как это обычно и принимается в грамматике), то сразу же видно, что слово *хороший* формально неоднородно:



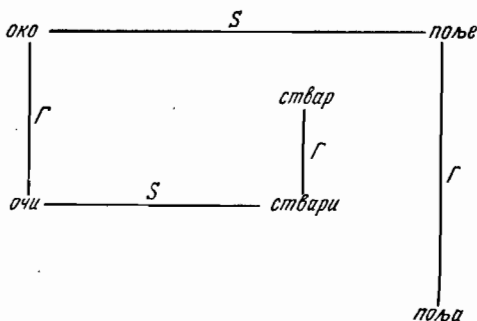
Во всех других славянских языках имеется параллелизм форм положительной и сравнительной степени (например, в чешском *hezký — hezčí*) и соответствующие слова формально однородны. Обычно это явление объясняется типологически как аналитичность образования склоняемых форм степеней сравнения в русском языке и это в известной мере <sup>6</sup> справедливо. Важно, однако, что критерий формальной однородности обнаруживает существенное отличие русского языка от всех остальных славянских языков. Важно также и то, что за вычетом четырех указанных слов все остальные прилагательные русского языка, сколько можно судить, формально однородны.

Другим интересным примером формально неоднородных слов являются недостаточные парадигмы типа *щец*, *дровец*, употребляемых только в этой форме. Легко видеть, что в этих примерах мы имеем дело также со словами с нулевой парадигматикой.

К нарушению формальной однородности приводит также наличие слов *singularia tantum*, а также *pluralia tantum*.

Типологически чрезвычайно интересны некоторые случаи формально неоднородных слов в сербском языке. Слово *око* в единственном числе эквивалентно слову *поле*, а во множественном числе *очи* эквивалентно слову *ствари*, ср. *чно*, *око*, *чно поле*, *чрни очи*, *ствари* при невозможности *\*чрни поля*.

Имеем схему:



Другие примеры из сербского языка с весьма интересной интерпретацией этого явления даются в работе Милки Ивич <sup>7</sup>.

<sup>6</sup> Все же остаются сомнения по поводу того, нужно ли считать *более красивый*, *самый красивый* формами одного слова или словосочетаниями. Тогда, вообще говоря, нужно было бы выделить столько степеней сравнения, сколько есть сочетаний типа *очень красивый*, *менее красивый*, *весьма красивый* и т. п. Сравни также проблему болгарских форм сравнительной степени (стр. 116), где рассмотрение форм с *по* (соответствующих *более*) приводило к неправильности.

<sup>7</sup> М. I v i ć. Relationship of gender and number in serbocroatian substantives. «International journal of slavic linguistics and poetics», VI, 1963.

Аналогичные примеры имеются в румынском языке, где выделяют специальный «обоюдный» (ambigen) или «средний» (neutru) род:

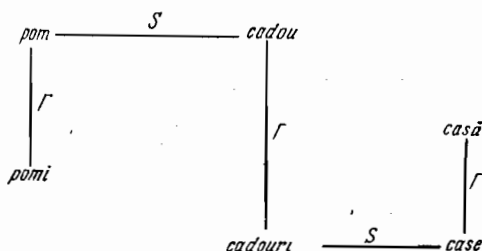
единственное число

un pom bun  
un cadou bun  
o casă bună

множественное число

pomi bun i  
 cadouri bune  
 case bune

Это дает схему:

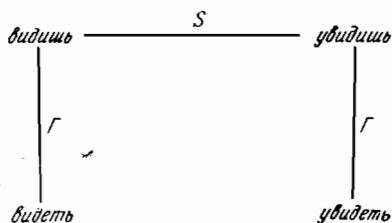


Здесь интересна структурная близость схем для двух генетически разных языков. Подобные явления имеют место и в других балканских языках, например, в некоторых диалектах албанского, хотя вряд ли можно считать их балканизмом (ср. отсутствие «обоюдного рода» в новогреческом, а с другой стороны, типологически подобные схемы в тохарском языке<sup>8</sup>).

Приведем еще один пример неоднородности. В § 35 мы будем говорить о причинах, по которым разные видовые формы, например, *видишь* и *увидишь* целесообразно считать принадлежащими к разным парадигмам. Очевидно, однако, что *видишь* и *увидишь*, *видит* и *увидит* суть словоформы, принадлежащие одному семейству. С другой стороны, словоформы *видеть* и *увидеть* принадлежат разным семействам, ср. *буду*, *начинаю*, *продолжаю*, *бросаю*, *прекращаю*, *прерываю*, *кончаю* *видеть* при невозможности \**буду увидеть*, \**начинаю увидеть*, \**кончаю увидеть* и т. п. К разным семействам принадлежат и формы *видим* и *увидим*, ср. *давай увидим* при невозможности \**давай видим*, по-видимому, и *вижу*, и *уввижу*, ср. *дай увижу* (или более обычное *дай посмотрю* при невозможности

<sup>8</sup> См.: Вяч. Вс. Иванов [замечания] «Тохарские языки». М., 1959, стр. 24; ср. также: A. Rosetti. Sur la catégorie du neutre. «Proceedings of the ninth International Congress of Linguists», The Hague, 1964, стр. 774—781.

\*дай вижу (или \*дай смотрю). Получаем следующую схему, свидетельствующую о нарушении однородности:



Распространим теперь понятие однородности и на случай одностороннего замещения. Это необходимо сделать потому, что нарушение принципа формальной однородности очень часто не свидетельствуют ни о каких специфических особенностях языка (как это было, например, в случае с русскими прилагательными или сербскими существительными, меняющими род), а связаны с тем, что в игру вступает омонимия.

Рассмотрим, например, следующий очень характерный фрагмент русского языка.

Фразы: *юноша стоит, стол стоит, вижу юношу, вижу стол, любуюсь юношей, любуюсь столом*

Парадигмы:  $\Gamma_1 = \{\text{юноша, юношу, юношей}\}$ ,  $\Gamma_2 = \{\text{стол, столом}\}$ ,  $\Gamma_3 = \{\text{любуюсь}\}$ ,  $\Gamma_4 = \{\text{вижу}\}$ ,  $\Gamma_5 = \{\text{стоит}\}$ .

Отношения замещения:

*юношу*  $\Rightarrow$  *стол*; *юноша*  $\Rightarrow$  *стол*; *столом*  $\Rightarrow$  *юношей*;

*юношей*  $\Rightarrow$  *столом*

других замещений нет.

Семейства:

$S_1 = \{\text{любуюсь}\}$	}	начальные
$S_2 = \{\text{стоит}\}$		
$S_3 = \{\text{юноша}\}$		
$S_4 = \{\text{юношу}\}$		
$S_5 = \{\text{юношей, столом}\}$		
$S_6 = \{\text{стол}\}$		

Элементарные грамматические категории:

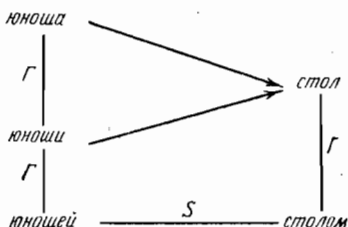
$G_1 = \{\text{юноша, стол}\}$ , порождена семейством  $S_3$

$G_2 = \{\text{юношу, стол}\}$ , порождена семейством  $S_4$

$G_3 = \{\text{юношей, столом}\}$ , порождена семейством  $S_5$

Остальные элементарные грамматические категории одноэлементны.

Полученные отношения иллюстрируются схемой:



Здесь нарушено условие формальной однородности:

$\Gamma(\text{стол}) \cap S(\text{юношей}) \neq \emptyset$ , а  $S(\text{стол}) \cap \Gamma(\text{юношей}) = \emptyset$ .

Между тем отношения между словами здесь вполне регулярны. Введем новые определения<sup>9</sup>:

0. 5. 20 Назовем полностью однородным всякий формально однородный язык с ненулевой парадигматикой.

0. 5. 21. Язык называется ограниченно однородным, если в нем для любых  $x$  и  $y$  из существования  $w$  такого, что  $w \in \Gamma(x)$  и  $w$  непосредственно связано с  $y$ , следует существование слова  $z$  такого, что  $z \in \Gamma(y)$  и  $z$  непосредственно связано с  $x$ . Так же, как и в случае формальной однородности, мы локализуем теперь введенные понятия, т. е. отнесем их не к языку в целом, а к отдельным словам.

0. 5. 22. Слово  $x$  называется полностью однородным, если оно формально однородно и имеет ненулевую парадигматику.

0. 5. 23. Слово  $x$  называется ограниченно однородным, если для любых слов  $y$  и  $z$  таких, что  $y \in \Gamma(x)$  и  $z$  непосредственно связано с  $x$ , найдется слово  $w$  такое, что  $w \in \Gamma(z)$  и  $w$  непосредственно связано с  $y$ .

В приведенном примере ни одно из слов не является полностью однородным. Зато каждое из слов, как легко проверить, ограниченно однородно. Отсюда, очевидно, следует, что приведенный фрагмент (который, разумеется, правилен) является ограниченно однородным. Мы видели на примерах, что не всякий формально однородный язык ограниченно однороден. Однако имеют место следующие утверждения:

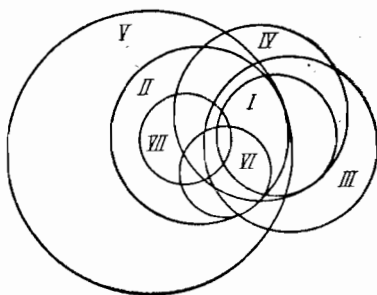
Т. 5. 6. Всякий полностью однородный язык ограниченно однороден.

Т. 5. 7. Всякий полностью однородный язык полон.

<sup>9</sup> Более подробно соответствующую систему определений см.: И. И. Ревзин. Распространение теоретико-множественной модели на случай грамматической омонимии. «Научно-техническая информация», 1965, № 3.

Отношения между введенными типами языков иллюстрируются следующей схемой:

- I — полностью однородные языки
- II — формально однородные языки
- III — языки с ненулевой парадигматикой
- IV — ограниченно однородные языки
- V — правильные языки
- VI — аморфные языки
- VII — предельно грамматикализованные языки.



Для того чтобы выяснить место на этой схеме других понятий этого параграфа, приведем еще некоторые факты.

**Т. 5. 8.** Всякий полный и ограниченно однородный язык является вполне правильным.

Из этого утверждения непосредственно следует аналог теоремы 5.4 для случая одностороннего замещения.

**Т. 5. 9.** Всякий полностью однородный язык является вполне правильным.

Смысл понятия полностью однородного языка в следующем. Модель формально однородного языка с лингвистической точки зрения неудобна тем, что она включает в себя некоторые тривиальные случаи. Например, имеет место следующая очевидная теорема.

**Т. 5. 10.** Всякий аморфный язык и всякий предельно грамматикализованный язык являются формально однородными.

Тем самым это понятие моделирует не только агглютинативные, но и изолирующие языки, хотя с лингвистической точки зрения эти группы языков весьма различны.

Интересна следующая теорема, ограничивающая применимость понятия полной однородности к аморфному языку.

**Т. 5. 11.** Аморфный язык полностью однороден тогда и только тогда, когда для любых  $x$  и  $y$  из  $x \Rightarrow y$  следует  $x \in S(y)$ .

Ближе всего к модели полностью однородного языка, по-видимому, агглютинативные языки<sup>10</sup>, в то время как ограниченная однородность характерна для флективных языков.

Так, если мы рассмотрим систему имени славянских языков, то подавляющее число слов здесь ограничено однородно (некоторые примеры отклонения от ограниченной однородности иллюстрируются примерами, приведенными выше в связи с отклонением от формальной однородности).

<sup>10</sup> Между прочим, Маркус предложил определить агглютинативные языки следующим способом: язык называется агглютинативным, если для

В типологических целях интересно однако рассмотреть отклонения от ограниченной однородности на некоторых важных фрагментах. Возьмем, например, все фразы без придаточных и без конструкций типа *одни из, некоторые из* и т. п. (ср. § 26), причем пусть для простоты в них существительные выступают только в роли подлежащих. На этом фрагменте разные славянские языки дают разные соотношения между  $D$ -классами слов в единственном и множественном числе, в связи с чем возникает отсутствие ограниченной однородности. Известно, что в единственном числе во всех славянских языках все существительные распадаются на три  $D$ -класса, соответствующих мужскому, женскому и среднему.

Что касается распределения  $D$ -классов во множественном числе, то славянские языки распадаются на три группы:

а) сербский тип:  $D$ -классы во множественном числе полностью соответствуют  $D$ -классам единственного числа, что дает ограниченную однородность (за исключением случая, рассмотренного на стр. 121), ср. *млади јелени млада дрвета, младе жене*;

б) русский тип: во множественном числе всего один  $D$ -класс, поскольку на этом фрагменте любые два слова во множественном числе непосредственно связаны *молодые олени, деревья, жены*.

в) чешско-польский тип, выделяется во множественном числе несколько  $D$ -классов, но они не соответствуют  $D$ -классам, выделенным в единственном числе. Это связано с выражением категории личности (*męska osobowość*) в польском языке *dobrzy uczniowie, dobre zeszyty, żony, dzieci* и категории одушевленности в чешском; *noví žáci, nové sešity, tužky, nova pera*.

При этом чешский язык тяготеет по поведению других  $D$ -классов к сербскому типу, а польский к русскому (более подробно см. в § 47, посвященном категории рода в славянских языках).

Отношения между соответствующими  $D$ -классами в единственном и множественном числе выражаются таблицей на стр. 127.

Здесь классы  $D_1$  соответствуют личному роду (в польском) и одушевленному (в чешском),  $D_2$  — неличному (включая и неодушевленный),  $D_3$  — женскому роду,  $D_4$  — среднему роду.

С типологической точки зрения интересно, что все славянские языки (каждый из которых относится к одному из приведенных четырех типов) можно сравнивать, исходя из того, какова для рассматриваемого фрагмента степень устранимости неоднородности. Из приведенной таблицы сразу же следует, что устранимость неоднородности затрудняется по мере движения от сербского типа к русскому. В самом деле: в чешском языке для достижения

---

любых  $x$  и  $y$  из  $x \Rightarrow y$  следует  $x \in S(y)$ , т. е. это как раз то требование, которое фигурирует в теореме 5.11 (см.: S. M a r c u s. Typologie des langues et modèles logiques. «Acta Mathematica Academiae Scientiarum Hungaricae», t. XIV, fasc. 3—4, 1963, стр. 279).

Однако одностороннее замещение все же имеется в агглютинативных языках: это все случаи употребления местоимений.

Язык	Число	Д <sub>1</sub>	Д <sub>2</sub>	Д <sub>3</sub>	Д <sub>4</sub>
Сербский	ед.	_____		_____	_____
	мн.	_____		_____	_____
Чешский	ед.	_____		_____	_____
	мн.	_____	_____	_____	_____
Польский	ед.	_____		_____	_____
	мн.	_____	_____	_____	_____
Русский	ед.	_____		_____	_____
	мн.	_____	_____	_____	_____

однородности достаточно исключить все предметные (неодушевленные) существительные мужского рода или же все существительные женского рода — в польском нужно устранить и те и другие сразу, а в русском нужно устранить, например, все существительные мужского и женского или мужского и среднего родов.

### § 35. ВОПРОС О РАЗГРАНИЧЕНИИ ИНДИВИДУАЛЬНЫХ ПАРАДИГМ. ЯЗЫК С ХОРОШО ЗАДАНЫМИ ПАРАДИГМАМИ

Остановимся на общих условиях отнесения двух слов к одной парадигме, исходя из основного понятия данной главы, а именно понятия однородности.

По-видимому, распределение слов по парадигмам целесообразно задавать так, чтобы

а) все слова, входящие в одну парадигму, несли одинаковую экстралингвистическую информацию, в частности, имели тот же самый денотат;

б) объединение слов в парадигмы давало бы возможность экономного, системного описания фактов языка.

Первое требование удовлетворяется тем, что слова, входящие в одну парадигму, имеют общую корневую морфему. Это соответствует лингвистической интерпретации парадигм как системы форм одного слова. Единственное отклонение, допускаемое здесь, это объединение в одну парадигму так называемых супплетивных форм, т. е. алломорфов одной и той же лексической морфемы (например, *есть* и *быть*, *хорошо* и *лучше* и т. п.). Итак, слова одной парадигмы должны иметь одну и ту же лексическую морфему.

Однако это необходимое требование не является достаточным. В самом деле, нецелесообразно, с точки зрения лингвистической интерпретации включать в одну парадигму любую пару из ряда слов: *человечек*, *человечный*, *человеческий*, *человечество*, *человечина*, *человечище*, *очеловечить*, хотя в них одна и та же морфема с алломорфами *человек/человеч*. Конечно, интуитивно разница между этими словами и невозможность объединения их в одну парадигму ясна, хотя бы потому, что многие из них принадлежат к разным частям речи. Но «часть речи» во всех наших моделях (как раздел, тип и гипертип) определялась через понятие парадигмы, поэтому это соображение не может использоваться при распределении слов с одинаковой лексической морфемой по парадигмам.

Здесь возникают проблемы, по поводу которых в лингвистике до сих пор нет единства. Должны ли мы считать, что *видеть* и *увидеть*, соответственно *вижу* и *увиджу*, *видишь* и *увидишь*, принадлежат одной или разным па-



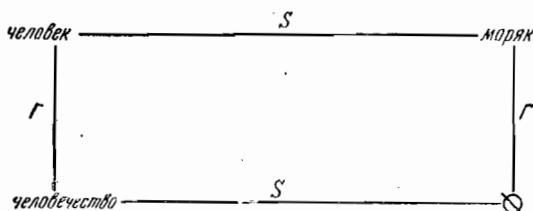
парадигмам? В традиционных терминах эту проблему можно поставить следующим образом: относится ли образование видовых форм к словообразованию (две разные парадигмы) или к формообразованию<sup>11</sup>? Попробуем подойти к решению этого вопроса, исходя из понятий, связанных с нашим вторым критерием, а именно критерием, отражающим системность описания языка.

Рассмотрим наш пример с видовыми формами. Если считать, что *видеть* и *увидеть*, *вижу* и *увиджу* и т. п. принадлежат к одной парадигме, то мы получаем нарушение принципа простоты во всех парадигмах, ибо *видишь* и *увидишь*, *видит* и *увидит*, как мы указывали, принадлежат одному семейству.

Поскольку разделение парадигм, т. е. соглашение о том, что две видовые формы суть разные слова, приводит к простоте во всех глагольных парадигмах, то мы и примем такое разбиение на парадигмы (тем более, что оно даст нам в § 46 возможность формального определения категории вида).

Итак, при объединении слов в парадигмы для таких языков, как славянские (как мы видели, в агглютинативных языках дело обстоит сложнее), целесообразно соблюдать требование простоты.

Это сразу же заставляет нас отнести к разным парадигмам *человек* и *человечек*, *человечный* и *человеческий* и т. п. Остается вопрос, почему к разным парадигмам следует отнести например, *человек* и *человечество*. Здесь в силу вступает критерий однородности. В самом деле *человек* и *морях* относятся к одному семейству. Если *человек* и *человечество* относятся к одной парадигме, то получаем схему, свидетельствующую о неоднородности:



Подобное же рассуждение можно провести и для остальных слов.

Конечно, этих двух требований недостаточно, ибо им можно удовлетворить тривиальным способом, превратив каждое слово в аморфное, т. е. оставив в каждой парадигме ровно по одному слову.

Поэтому мы введем еще третье требование.

0.5.25. Будем говорить, что в языке хорошо заданы парадигмы, если любое их укрупнение приводит к увеличению числа простых и неоднородных слов.

Следует особо рассмотреть вопрос об образовании множественного числа существительных. В примерах из сербского и румынского языка в § 34 неоднородность возникала за счет того, что формы единственного и множественного числа существительных объединены в одну парадигму, то же относится к сравнению четырех славянских языков по критерию степени их неоднородности.

Точка зрения, по которой образование множественного числа есть образование нового слова, а не формы того же слова, неоднократно высказывалась в славистике<sup>12</sup>.

<sup>11</sup> О разных подходах к этой проблеме см.: В. В. Виноградов. Современный русский язык. М., 1938, стр. 380—390.

<sup>12</sup> См., например: П. С. Кузнецов. Развитие индоевропейского склонения в общеславянском языке. М., 1958, стр. 7.

По-видимому, однако, целесообразно строить модель, одинаково описывающую язык и его наиболее характерные подязыки. Так, по-видимому, для славянских языков верно то, что исключение некоторых слов вроде сербского *очи*, а также *pluralia tantum* делает языком с хорошо заданными парадигмами язык, в котором объединены формы единственного и множественного числа.

В общем представляется более целесообразным такое решение вопроса, при котором для большинства слов можно сохранить более крупные парадигмы.

### § 36. ОТНОШЕНИЕ ТРАНСФОРМИРУЕМОСТИ

До сих пор мы занимались отношением между отдельными словами. Теперь мы будем исследовать типы языков, возникающие в связи с рассмотрением определенного отношения между фразами, а именно отношения трансформируемости. Прежде всего введем следующее определение.

0.5.26. Фраза  $f = x_1 \dots x_n$  находится в отношении  $\Gamma$ -связи, или  $\Gamma$ -связана, с фразой  $g = y_1 \dots y_m$ , если для каждого неаморфного слова  $x_i$  в  $f$  найдется  $y_j$  в  $g$  такое, что  $y_j \in \Gamma(x_i)$  и для каждого неаморфного  $y_i$  из  $g$  найдется  $x_j$  в  $f$  такое, что  $x_j \in \Gamma(y_i)$ . Фраза  $f = x_1 \dots x_n$  находится в отношении трансформируемости с фразой  $g = y_1 \dots y_m$ , если 1) они  $\Gamma$ -связаны, и 2) существуют  $\Gamma$ -связанные и тождественные по содержанию реальные фразы  $f' = u_1 \dots u_n$  и  $g' = w_1 \dots w_m$  такие, что  $u_i \in S(x_i)$  и  $w_j \in S(y_j)$ .

По поводу этого определения необходимо сделать следующие замечания.

а) По-видимому, для решения практических задач, лучше использовать несколько другое определение, в котором, в пункте 2 «семейства» заменяются на  $\mathcal{D}$ -классы (это, как мы видели, вообще необходимо во всех тех случаях, когда мы не можем отвлечься от фактов омонимии).

б) При интерпретации определения  $\Gamma$ -связанности следует понимать парадигму уже не как узкую парадигму склонения или спряжения, а более широко как совокупность форм с той же основой. Мы иногда будем называть такую парадигму семантической (ср. гл. VIII). Например, мы будем при этом считать, что отглагольное существительное *открытие* попадает в ту же парадигму, что и глагол *открывает*, причастие *открывающий*, деепричастие *открывая* и т. п. Мы употребляем тот же формальный термин, ибо можно считать, что парадигмы каждый раз задаются в зависимости от целей исследования, в то же время использование ряда понятий, введенных ранее, например, понятия однородного языка, при этом существенно облегчается (не надо вводить никаких новых терминов). Конечно, интерпретации, данные ранее при другой трактовке парадигмы, здесь уже не действительны (это относится, например, к моделированию «части речи» через парадигму).

в) Определение трансформируемости использует отношение тождественности смысла двух реальных фраз. Это отношение (а значит и отношение трансформируемости) рефлексивно, симметрично, но не транзитивно, что оправдано из-за существования двусмысленных фраз. Здесь трансформация прямо отражает семантику.

В § 18 указывалось, что задание смыслового тождества вполне допустимо в теории моделей. Что же касается понятия трансформации, то оно, как стало очевидно подавляющему большинству лиц, работающих в этой области, часто покоится на семантике, а именно любой исследователь, вырабатывающий формальные условия для того, чтобы две фразы  $A$  и  $B$  трансформировались одна в другую, сам опирается на то, что эти фразы тождественны по смыслу или же что фразы  $f$  и  $g$  различаются лишь некоторой семантической константой (например, отрицанием). Мы могли бы ввести последнее требование<sup>13</sup> в определение, но предпочли не усложнять его.

Нам важно другое, а именно сформулировать понятие трансформации независимо от грамматики, порождающей множество фраз языка. Это вызвано тем соображением, что смысловое тождество двух фраз вовсе не должно вытекать из способа их порождения (как это принимается в ряде концепций трансформационной грамматики). Мы хотим, чтобы отношения, связанные со смысловым тождеством и вообще со смыслом, устанавливались для готовых фраз. Нам представляется, что именно такое понимание вытекает из всего опыта традиционной грамматики (более полное обоснование см. в § 51).

В дальнейших формальных исследованиях мы пользуемся только свойством  $T$ -связанности. Таким образом, требование тождественности смысла скорее относится к лингвистической интерпретации модели, чем к ее внутренней структуре. Несмотря на это, мы считаем, однако, что явное введение этого требования чрезвычайно важно для развития лингвистической теории моделей.

### § 37. ЯЗЫК С $T$ -ОБУСЛОВЛЕННЫМИ ПАРАДИГМАМИ И ЯЗЫК С ПОЛНОЙ ТРАНСФОРМАЦИЕЙ

0. 5. 27. Некоторая парадигма  $\Gamma(x)$  называется  $T$ -обусловленной, если для любой фразы  $f$ , содержащей  $x$ , и для любого слова  $y$ , входящего в  $\Gamma(x)$ , найдется фраза  $g$ , содержащая  $y$  и находящаяся в отношении трансформируемости с  $f$ .

Имя в славянских языках устроено так, что парадигмы его не являются  $T$ -обусловленными. Например, возьмем фразу: *Плотник строит дом*. Для некоторых слов из парадигмы  $\Gamma$  (*плот-*

<sup>13</sup> Оно предложено проф. Хижем (Hiż) на симпозиуме по семиотике в Польше в сентябре 1966 г.

ник), например *плотником*, находятся трансформированные фразы, но для других (*плотнике, плотнику*) таких фраз нет (фраза *плотнику строит дом*, хотя и является  $\Gamma$ -связанной с нашей фразой, но имеет другой смысл).

Другое дело глагол в славянских языках. Возьмем например, фразу *я читаю*. Любая из форм, входящих в парадигму  $\Gamma$  (*читаю*), входит в некоторую фразу, связанную отношением трансформируемости с нашей фразой: *ты читаешь* и т. п. Это верно для любых цепочек, имеющих тот же  $S$ -образ.

<i>я читаю</i>	<i>я рисую</i>
<i>ты читаешь</i>	<i>ты рисуешь</i>
<i>он читает...</i>	<i>он рисует...</i>

Этот пример позволяет сформулировать еще одно понятие.

0. 5. 28. Фраза  $f = x_1 \dots x_n$  называется фразой с полной трансформацией, если для любых фраз  $f' = x'_1 \dots x'_n$ , где  $x'_i \in S(x_i)$ , и  $g = y_1 \dots y_m$ , находящейся в отношении трансформируемости с  $f$ , существует фраза  $g' = y'_1 \dots y'_m$ , находящаяся в отношении трансформируемости с  $f'$ , причем  $y'_i \in S(y_i)$  и если  $x_i \in \Gamma(y_j)$ , то  $x'_i \in \Gamma(y'_j)$  и обратно.

0. 5. 29. Языком с полной трансформацией (ЯПТ) называется язык, в котором каждая фраза есть фраза с полной трансформацией.

Примечание. При сравнении этих определений с теми, которые предложены в «Моделях языка» (стр. 145), легко заметить, что определение трансформации, предложенное там, почти целиком совпадает с определением полной трансформации, как оно сформулировано здесь. Преимущество новой системы определений, как нам представляется, в том, что полнота трансформации локализована как типологическое свойство некоторого подмножества всего множества фраз (то же намерение было у автора и тогда, но сформулированное в «Моделях языка» понятие языка с полной трансформацией при той же лингвистической интерпретации было с формальной точки зрения некорректным).

Как мы видели, свойства полной трансформации и  $T$ -обусловленности парадигм, хотя они с логической точки зрения и независимы как правило выступают вместе. Интересно, что имеет место следующая теорема:

Т. 5. 12. Язык с полной трансформацией, в котором все парадигмы  $T$ -обусловлены, формально однороден.

Отсюда сразу же следует, что подязык, содержащий только глаголы и личные местоимения, в именительном падеже однороден.

Свойство  $T$ -обусловленности парадигм приводит и к другому любопытному обстоятельству. Возьмем некоторую  $T$ -обусловленную парадигму  $\Gamma(x)$  и некоторую фразу  $f$ , содержащую  $x$ . Если эту фразу рассматривать как единое слово, то всю совокупность фраз, требуемых определением 5.27, можно интерпретировать как совокупность словоформ, относящихся к одной расширенной парадигме.

Примечание. Д. Уорт<sup>14</sup> предпринял интересную классификацию соответствующих расширенных парадигм. Он различает, в частности, два важных типа парадигм:

а) «простые линейные парадигмы»

русский

<i>я его видел</i>	<i>мы его видели</i>
<i>ты его видел</i>	<i>вы его видели</i>
<i>он его видел</i>	<i>они его видели</i>
<i>я ее видел...</i>	<i>мы ее видели...</i>

Подобные фразы можно уподобить агглютинативным формам, где каждый формант сигнализирует об определенной категории. Как мы уже видели, агглютинативность влечет за собой формальную однородность и поэтому формальная однородность соответствующего фрагмента не должна нас удивлять.

Д. Уорт, по-видимому, ошибочно причисляет к «простым линейным парадигмам» соответствующие польские формы: *widziałem go; widziałeś go* и т. п.

По поведению местоимения эти парадигмы действительно изоморфны русским. Однако при употреблении с наречием, вопросительными местоимениями и т. п. показатель лица может отделяться, ср. *czytałem, no książkęm czytał, czytałaś — czyś czytała, widzieliście — dobreście widzieli*. Это примыкание показателя лица к первому слову ритмической группы скорее напоминает статус кратких местоименных форм в следующем типе, выделенном Д. Уортом.

б) «простые нелинейные парадигмы»

сербско-хорватский

<i>я сам га видео</i>	<i>я сам је видео</i>
<i>ти си га видео</i>	<i>ти си је видео</i>
<i>он га је видео...</i>	<i>он ју је видео...</i>

Нельзя заменить порядок в третьей фразе *он је га видео*, сделав его аналогичным первым двум. Это уже отход от принципа агглютинативности, однако если рассматривать *сам га, си га, га је* как единые морфы, мы еще имеем агглютинативный принцип с вытекающей из него формальной однородностью. Аналогично, если считать, что *ју је* вместо *је је* единый морф.

Уорт считает, что для типологии в качестве расширенных парадигм целесообразно рассматривать и случаи типа

сербско-хорватский	польский	русский
<i>он је професор</i>	<i>он jest profesorem</i>	<i>он профессор</i>
<i>он је био професор</i>	<i>он byl profesorem</i>	<i>он был профессором</i>

Здесь действительно вскрывается типологическое различие между польским и сербско-хорватским, с одной стороны, где при трансформации заме-

<sup>14</sup> Автор имел возможность обсудить соответствующую проблематику с американским ученым во время пребывания последнего в Москве и пользуется случаем выразить ему благодарность. Взгляды Уорта изложены также в работе: Dean S. Worth. The role of transformations in the definition of syntagmas in Russian and other slavic languages. «American contributions to the 5. International Congress of Slavists, Sofia. 1963». The Hague, Preprint, 1963, стр. 64.

няется всего один элемент, и русским, с другой стороны (хотя с материальной точки зрения русский язык находится посредине между польским и сербско-хорватским). Однако здесь мы выходим за пределы фрагмента, ограниченного глаголами и относящимися к нему местоимениями, которые и интонационно несамостоятельны и примыкают к глаголу.

Расширение парадигмы за счет объединения в один элемент целой группы слов проводится нами в сущности по следующему принципу. Имеется одно слово  $x$ , имеющее лексическое значение, т. е. принадлежащее к полнзначной парадигме, и ряд слов  $y_1, y_2 \dots y_m, z_1, z_1$ , имеющих служебный характер. Объединенный сегмент относится нами к парадигме  $\Gamma(x)$  как одна из словоформ этой парадигмы. Итак, перед нами возможность двоякой трактовки некоторой последовательности:

$$y_1 \dots y_m x z_1 \dots z_1$$

Такая ситуация в типологии языков встречается достаточно часто, ср. возможность двоякой трактовки элемента  $ga$  в эстонском как послелога или суффикса комитатива, или же аналогичное положение элемента  $nen(en)$ ,  $men(en)$ ,  $ben(en)$ , в казахском языке. Разная сегментация соответствующих фактов приводит в традиционной морфологической типологии к резко различной типологической оценке соответствующих языков.

Интересно поэтому проследить, влияет ли операция укрупнения сегмента на выделенные нами модельные свойства.

Что касается свойства простоты, то оно не сохраняется при укрупнении сегмента.

В этом одна из причин того, что агглютинативные формы, которые можно рассматривать как полученные в результате укрупнения из аморфных, часто ведут к нарушению простоты, так знаменательно происхождение эстонского суффикса комитатива  $ga$ , рассмотренного нами в связи с нарушением простоты («Модели языка», стр. 86) из послелога (ср. финский послелог  $kansa$ )<sup>15</sup>.

Перейдем теперь к понятию формальной однородности. При нашей трактовке операции укрупнения сохранение формальной однородности означает, что формально однородному слову  $x$  должна соответствовать фраза с полной трансформацией  $y_1 \dots y_m x z_1 \dots z_1$  (в самом деле, слову, входящему в то же семейство, должна соответствовать фраза с тем же  $S$ -образом, а слову, входящему в ту же парадигму, фраза,  $\Gamma$ -связанная с нашей). В этом свете приобретает особый смысл следующее утверждение:

<sup>15</sup> Об этом и аналогичных явлениях (разумеется, вне связи с понятием «простоты», но зато выдвигая на первый план «стандартность», формализуемую нами как формальную однородность) пишет Б. А. Серебренников («О причинах устойчивости агглютинативного строя». — ВЯ, 1963, № 1, стр. 49—54).

**Т. 5. 13.** Всякий простой и формально однородный язык является языком с полной трансформацией.

**Примечание.** В доказательстве требование простоты используется лишь со ссылкой на теорему 5.5. Если же взять язык, в котором всякое слово, эквивалентное аморфному, само аморфно (а это свойство выполняется в любом естественном языке), то тогда утверждение теоремы 5.13 действительно для любого формально однородного языка.

Из этой теоремы непосредственно следует, что операция укрупнения сегмента сохраняет свойство формальной однородности. Таким образом, это свойство (как и все связанные с ним) дает нам весьма глубокую характеристику языка, не зависящую от длины отрезка, рассматриваемого как слово.

**Примечание.** Отсюда еще не следует, что свойство формальной однородности инвариантно относительно обратной операции. Из теоремы 5.13 следует лишь, что если при некоторой сегментации на слова язык был формально однороден, то он останется таковым и при более крупных сегментах. Обратное, вообще говоря, неверно. Скорее дело обстоит так, что практически всякое укрупнение сегментов ведет к повышению формальной однородности. Связано это с тем, что отдельные слова менее ограничены в своих употреблениях, чем связанные морфы.

Расширение понятия парадигмы в связи с исследованием трансформации может быть произведено в другом направлении, а именно для исследования вопроса о том, насколько введенные в этом параграфе понятия применимы к таким отглагольным образованиям, как причастия, имена действия, имена деятеля и т. п.

Здесь полезно следующее понятие, введенное в несколько других терминах З. М. Волоцкой<sup>16</sup>. Пусть понятие парадигмы берется в самом широком смысле (как совокупность слов, имеющих общую лексическую морфему).

**0. 5. 30.** Слово  $x$  находится в отношении производности к  $y \in \Gamma(x)$ , если для любой фразы  $f$ , содержащей  $x$ , найдется фраза  $g$ , содержащая  $y$  и находящаяся в отношении трансформируемости к  $f$ .

Волоцкая приводит следующие примеры. Возьмем слова *сложный* и *сложность*. Рассмотрим фразу *сложный вопрос*. Для нее можно найти трансформированную фразу *сложность вопроса*. Аналогично для любых фраз, содержащих слово *сложный*. Отсюда следует, что слова *сложность* и *сложный* связаны отношением производности. Рассмотрим теперь слова *уравнение* и *уравнять*. Возьмем фразу *квадратное уравнение*. Не существует фразы, находящейся к ней в отношении трансформируемости и содержащей слово *уравнять*. Отсюда следует, что слова *уравнение* и *уравнять* не находятся в отношении производности.

<sup>16</sup> З. М. Волоцкая. Установление отношения производности между словами (опыт применения трансформационного метода). — ВЯ, № 3, 1960, стр. 100—107.

Надо сказать, что приведенная нами формулировка отношения производности по существу отличается от формулировки Волоцкой тем, что определенное нами отношение производности несимметрично, в то время как Волоцкая подчеркивала важность симметричности соответствующего отношения. Мы сделали это для того, чтобы подчеркнуть одну идею, содержащуюся в работе Волоцкой, но явно не выраженную. Для ясности рассмотрим приводимый ею пример. Слово *предложить* находится, как легко проверить, в отношении производности к слову *предложение*. Обратное неверно, так как для фразы *сложносочиненное предложение* нельзя найти трансформированной фразы, содержащей слово *предложить*. Будем теперь считать, что во всех контекстах, не поддающихся обратной трансформации, мы имеем дело с другим словом. В результате такой операции слова *предложить* и *предложение* окажутся взаимно-производными и отношение производности станет симметричным.

Отношение взаимной производности важно с той точки зрения, что расширение некоторой парадигмы за счет слов, находящихся в отношении производности со словами данной парадигмы, не меняет свойства *T*-обусловленности.

С типологической точки зрения очень интересно проследить, что же происходит при этом с другим выделенным нами свойством, а именно свойством полной трансформации. В каждом языке можно выделить множество фраз с *T*-обусловленными парадигмами и множество фраз с полной трансформацией. Важным типологическим критерием может служить разный объем этой части в различных языках. Так, в польском языке почти от каждого глагола можно образовать имя действия с суффиксом *-anie/-enie*. Если при этом глагол имел при себе показатель возвратности *się*, то соответствующий показатель остается и при отглагольном имени<sup>17</sup>, например, *Co się dzieje podczas uczenia się drugiego języka i posługowania się nim?*

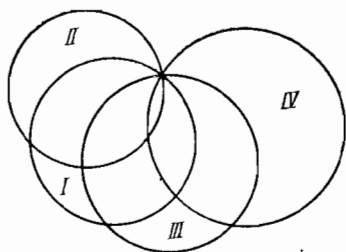
Таким образом, в любой фразе, в какой нам ни встретится какая-либо личная форма глагола *uczyć*, ее можно трансформировать во фразу с именем *uczenie*. Если мы при этом считаем *się* отдельным словом, то соответствующая трансформация возможна и для такого контекста, т. е. мы имеем дело с *T*-обусловленной парадигмой (ср. невозможность трансформации при трактовке русского *-ся* в *обучаться* как отдельного слова). Если же считать *się* частью глагольной формы, то любая фраза, имеющая тот же *S*-образ, что и фраза *uczyć się drugiego języka*, есть фраза с полной трансформацией.

<sup>17</sup> О возможности на этом основании включать имя действия в польском языке в парадигму глагола см.: J. Dąborski. Charakteristika polštiny v rogovnání s češtinou. «Studie z filologie polskiej i słowiańskiej», 5. Warszawa, 1965, стр. 265.



## § 38. ОБЩАЯ ОЦЕНКА ТЕОРЕТИКО-МНОЖЕСТВЕННОЙ ТИПОЛОГИИ

Интересно рассмотреть множество языков, обладающих выделенными нами типологическими характеристиками. На этом множестве имеют место отношения включения, показанные на стр. 125, а также на схеме (взяты лишь наиболее важные характеристики).



- I — формально однородные языки
- II — языки с *T*-обусловленными парадигмами
- III — языки с полной трансформацией
- IV — простые языки

Схемы показывают центральное положение формально однородных языков, находящихся на пересечении большинства других свойств. Это особенно интересно в связи с выявленной в § 37 независимостью этого свойства от операции укрупнения.

Из многочисленных примеров, приведенных в § 34, вытекает важность этого свойства для таких флективных языков, как славянские. Однако мы видели, что оно имеет место и для языков других типов, являясь, по-видимому, языковой универсалией.

Это не означает, что построенная теоретико-множественная типология не имеет ценности для сравнения языков. Прежде всего мы видели, что разные языки, стремясь в «пределе» к формальной однородности, по разному отклоняются от нее на разных правильных фрагментах, причем это дает возможность сравнить между собой такие близко родственные языки, как славянские.

Наряду с понятием формальной однородности, являющимся, как мы видели, центральным, мы получили ряд других полезных критериев для случаев, когда свойство однородности заведомо не выполняется.

Кроме того, в нашей типологии нашлось четкое место аморфным языкам (формально однородные, но, вообще говоря, не полностью однородные языки), агглютинативным языкам (полностью однородные языки) и флективным языкам (ограниченно однородные, простые) — характеристики, присущие этим языкам на основных их фрагментах. Хотя теоретико-множественная типология отражает лишь некоторые идеи традиционной типологии, основные критерии последней отражаются в ней достаточно полно.

## МОДЕЛИРОВАНИЕ ГРАММАТИЧЕСКИХ КАТЕГОРИЙ СЛАВЯНСКИХ ЯЗЫКОВ

### § 39. НЕКОТОРЫЕ ПРЕДВАРИТЕЛЬНЫЕ ОПРЕДЕЛЕНИЯ

Особенностью славянских, как и вообще флективных языков, является то, что отдельные грамматические категории в них не получают непосредственного выражения, а поэтому не могут получить простого определения.<sup>1</sup>

Поэтому для них приходится строить достаточно сложные модели.

Модели, которые будут строиться в данной главе, должны дать ответ на следующие вопросы:

- а) что значит, что два произвольных слова относятся к одной и той же конкретной категории (падежа, рода и т. п.);
- б) сколько таких категорий в данном (славянском) языке определяет данная категориальная функция;
- в) какие категории в разных частях речи имеют одинаковую структуру.

Мы увидим, что ответ на первый вопрос дает возможность однозначно решить и остальные две задачи. Однако, как и в главе IV, мы могли выделять разные гипертипы, но не могли установить формально, какие слова являются, например, существительными, а какие — прилагательными, так и здесь мы сможем ответить на вопрос, принадлежат ли два слова одному падежу (или роду), но не сможем сказать, к какому конкретному падежу (или роду) принадлежат данные слова. Говоря о именительном, родительном и других падежах или о женском, мужском и других родах, мы будем иметь в виду уже интерпретацию получившихся формальных разбиений. Вопрос о семантических моделях, определяющих и принадлежность к данной конкретной категории, частично рассматривается в главе VIII.

Решение поставленной нами здесь задачи существенно облегчается тем, что, как мы указывали в § 16, в теории лингвистических моделей грамматические категории рассматриваются как множества слов, получающиеся путем применения к словарю некоторой категориальной функции. Мы и будем теперь строить соответствующие функции, но предварительно введем некоторые определения:

**0.6.1.** Категория  $x_1$  называется синтагматической, если для любого  $x$  из  $x \in x_1$  следует  $S(x) \subseteq x_1$ .

**0.6.2.** Категория  $x_1$  называется парадигматической, если для любого  $x$  из  $x \in x_1$  следует  $\Gamma(x) \subseteq x_1$ .

Таким образом, всякая парадигматическая категория может быть представлена как объединение парадигм, а всякая синтагматическая категория как объединение семейств.

Рассмотренные два критерия классификации категорий, вообще говоря, независимы. Категория может быть одновременно синтагматической и парадигматической, но это своего рода редкий случай (таковы классы управления глагола). Имеет место, например, следующее почти очевидное утверждение:

**Т.6.1.** Пусть категория  $x_1$  является одновременно парадигматической и синтагматической и пусть  $x \in x_1$ . Тогда  $Ra(x) \subseteq x_1$ .

В рассматриваемых нами дальше моделях каждая непарадигматическая категория обязательно является синтагматической, т. е. объединением семейств.

Следующее понятие помогает связать понятие категории и понятие формально однородного языка.

**0.6.3.** Категория  $x_1$  равномерно распределена по парадигмам, если для любого  $y \in Ra(x)$  из  $\Gamma(x) \cap x_1 \neq \emptyset$  следует  $\Gamma(y) \cap x_1 \neq \emptyset$ .

**0.6.4.** Категория  $x_1$  равномерно распределена по семействам, если для любого  $y \in Ra(x)$  из  $S(x) \cap x_1 \neq \emptyset$  следует  $S(y) \cap x_1 \neq \emptyset$ .

Имеет место следующая теорема, связывающая эти понятия с формальной однородностью:

**Т.6.2.** Если язык формально однороден, то в нем все синтагматические категории равномерно распределены по парадигмам, а все парадигматические категории равномерно распределены по семействам.

До сих пор мы рассматривали свойства однородности, связанные с объединением в семейства и парадигмы. Теперь мы рассмотрим некоторые свойства однородности, связанные с объединением в категории. Для этого введем следующие определения.

**0.6.5.** Назовем категориальный пучок  $\Pi = (\pi, \sigma, x_1, \dots, x_n)$ , где  $\pi$  — парадигматическая,  $\sigma$  — синтагматическая, а  $x_1, \dots, x_n$  — произвольные категории, отмеченным, если имеется хотя бы одно такое слово  $x$ , что  $x \in \Pi$ .

**0.6.6.** Назовем язык категориально однородным, если в нем для любых трех отмеченных категориальных пучков  $\Pi_1 = (\pi_1, \sigma_1, x_1, \dots, x_n)$ ,  $\Pi_2 = (\pi_2, \sigma_2, x_1, \dots, x_n)$  и  $\Pi_3 = (\pi_3, \sigma_3, x_1, \dots, x_n)$  категориальный пучок  $\Pi_4 = (\pi_2, \sigma_2, x_1, \dots, x_n)$  будет отмеченным.

**Примечание.** Смысл понятия категориальной однородности в том, что если мы будем строить дерево логических возможностей (ср. § 10) для множества словоформ, используя категории как дифференциальные при-

знаки, то мы не встретим ни одной пустой клетки. Высокая степень системности грамматических объектов (особенно по сравнению с другими совокупностями объектов, изучаемыми в лингвистике) делает требование категориальной однородности вполне оправданным.

До сих пор категории задавались извне. Теперь мы переходим на другую точку зрения. Пусть язык категориально пуст (ср. определение 3.37), иначе говоря, задан как совокупность слов (словарь  $V$ ), множество фраз  $\theta$  и разбиение словаря на парадигмы (другие исходные объекты здесь не понадобятся), и ничего не известно о принадлежности слов к категориям (т. е. используются только понятия, определенные при помощи указанных здесь исходных).

Мы хотим теперь в этих терминах определить аналоги грамматических категорий и в первую очередь аналоги категорий падежа и рода.

#### § 40. МОДЕЛИ ФУНКЦИИ ПАДЕЖА СУЩЕСТВИТЕЛЬНЫХ. ПАДЕЖИ РУССКОГО ЯЗЫКА

Моделирование функции падежа было начато А. Н. Колмогоровым и В. А. Успенским<sup>1</sup>. Предложенная ими модель содержала, однако, такие неопределяемые в рамках теории лингвистических моделей термины как «состояние», «предмет» и т. п. Попытка изложить ту же идею в терминах теории моделей была принята автором<sup>2</sup> и уточнена в работе Маркуса<sup>3</sup>.

Суть всех этих устремлений состоит в том, чтобы определять падеж через определенный класс контекстов. Мы отчасти отразим эту идею в следующем предварительном определении.

0.\* 6. 7. Пусть у нас имеется некоторое фиксированное множество контекстов, которые называются падежеобразующими, и пусть множество падежеобразующих контекстов распадается на ряд равноемких классов. Тогда падежом называется совокупность слов, допускаемых одним из таких классов.

Это еще не определение, поскольку не определено исходное понятие падежеобразующих контекстов. Это некоторая схема определения, которая может наполняться тем или иным содержанием.

Так, Маркус в цитированной работе предложил определение падежа, исходящее из всего множества контекстов.

<sup>1</sup> См.: В. А. Успенский. К определению падежа по Колмогорову. «Бюллетень объединения по проблемам машинного перевода». М., 1957, № 5, стр. 11—18.

<sup>2</sup> И. И. Ревзин. Модели языка. М., 1962, стр. 68—69.

<sup>3</sup> S. Marcus. Modèles mathématiques pour la catégorie grammaticale du cas. «Revue roumaine de mathématiques pures et appliquées», t. VII, 1963, № 4, стр. 585—610.

Иначе говоря, по Маркусу падеж определяется следующим образом.

0. 6. 7'. П а д е ж о м называется совокупность слов, которая допускается одним из классов равноемких контекстов.

П р и м е ч а н и е. Единственное отличие нашего определения от формулировки Маркуса состоит в том, что он называет падежом класс контекстов, а о словах говорит, что они допускаются данным падежом. Но такое словоупотребление, идущее еще от модели Колмогорова—Успенского, противоречит лингвистическому употреблению термина «падеж».

Определение 6. 7' действительно вполне удобно, если рассматривать, например, следующие классы контекстов в русском языке:

Контекст	Допускаемое слово
1. (Λ, идет) (Λ, гуляет) (Λ, стоит) (Λ, подвергается обработке)	мальчик, девочка, окно, метро, стол, чай, шоколад, народ
2. (Λ, идут) (Λ, гуляют) (Λ, стоят) (Λ, подвергаются обработке)	мальчики, девочки, окна, столы, народы
3. (вот, Λ) (это, Λ)	мальчики, девочки, окна, столы, народы, мальчик, девочка, окно, стол, народ
4. (особенность, Λ) (удаляюсь от, Λ) (возникают из, Λ)	мальчика, девочки, окна, метро, стола, чая, шоколада, народа, мальчиков, девочек, окон, столов
5. (мало, Λ) (кило, Λ)	народу, чаю, шоколаду, народа, чая, шоколада
6. (принадлежность к, Λ) (присяга на верность, Λ)	мальчику, девочке, народу, шоколаду, чаю, народам, мальчикам, девочкам
7. (вижу, Λ) (стреляю в, Λ)	мальчика, девочку, окно, метро, стол, чай, шоколад, народ, мальчиков, девочек, окна, столы, народы
8. (не вижу, Λ) (не могу переделать, Λ)	мальчика, девочку, девочки, окно, метро, стол, окна, стола, шоколада, мальчиков, девочек, окна, столы, столов
9. (доволен, Λ) (сижу с, Λ)	мальчиком, девочкой, окном, столом, метро, чаем, шоколадом, народом, мальчиками, народами, девочками
10. (гуляю в, Λ) (нахожусь в, Λ)	мальчике, девочке, лесу, метро, поле, шоколаде, мальчиках, народах

11. (говору о, А) мальчике, девочке, лесе, метро, поле, шоколаде, мальчиках, народах
12. (дам детям, А) мальчика, девочку, окно, стол, шоколад, шоколада, мальчиков, девочек, окна, столы

Легко видеть, что, во-первых, нами отобраны контексты, однозначно репрезентующие «школьные» падежи русского языка. Это контексты 1 (или 2, или 3) — именительный падеж, 4 — родительный падеж, 6 — дательный падеж, 7 — винительный падеж, 9 — творительный падеж и 11 — предложный падеж. Во-вторых, отобраны контексты, уже привлекавшие к себе внимания ряда грамматистов в связи с проблемой падежа. Это контексты 5 — второй родительный, или «количественно отделительный» падеж, контекст 10 — второй предложный, или «местный», и наконец, контексты 8 и 12, которые обычно трактуются как допускающие формы двух падежей. Таким образом, мы представили все основные контекстные классы, влияющие на выбор падежной формы — и то, что каждый из таких классов, по определению 6.7', есть отдельный падеж, скорее следует рассматривать не как недостаток, а как достоинство модели.

Отметим следующие особенности данного определения:

- 1) если некоторое слово входит в данный падеж, то любое слово, входящее в то же семейство, входит в данный падеж, т. е. падеж категория синтагматическая;
- 2) одна и та же словоформа может допускаться одновременно несколькими падежеобразующими классами.

Эти особенности вполне соответствуют принятой грамматической практике, и это также говорит в пользу обсуждаемой модели.

Недостаток данной модели, однако, в том, что она применима не ко всему множеству контекстов, а только к некоторым специальным контекстам типа рассмотренных выше. Так, если рассматривать все множество контекстов русского языка, то слова *мальчик*, *девочка*, *окно* будут входить в разные падежи, так как контексты (*маленький, идет*) (*маленькое, идет*), (*маленькая, идет*) принадлежат к разным контекстным классам и, стало быть, относятся к разным падежам.

Другой недостаток приведенного определения, указанный автору А. А. Зализняком, состоит в том, что некоторые контексты допускают очень много форм, например:

(несу, А) — допускает слова: *собаку, собаке, рукой, воду, воды, чая, чаю* и т. п.

(дает, А) — допускает слова: *человек, воды, воду, сестре, книгу, рукой* и т. п.

Это вновь увеличило бы число падежей, а главное, привело бы к возникновению падежей, противоречащих грамматической интуиции.

Итак, прежде всего нам нужно некоторой специальной процедурой ограничить множество рассматриваемых контекстов и определять падежи, исходя в первую очередь из них.

0. 6. 8. Назовем контекст *диагностическим* для типа  $T_1$ , или короче  $T_1$ -диагностическим, если он допускает любую парадигму из  $T_1$ .

0. 6. 9.  $T_1$ -диагностической фразой называется фраза  $fxg$ , где  $(f, g)$  есть  $T_1$ -диагностический контекст, а  $x \in T_1$ .

Т. 6. 3. Если язык  $S$ - $T$ -правилен, то для любого  $T$  множество всех  $T$ -диагностических фраз есть правильный фрагмент.

Итак, множество всех  $T$ -диагностических фраз представляет собой правильный фрагмент (а мы раньше условились пользоваться лишь такими фрагментами языка).

Вообще же некоторая совокупность  $T$ -диагностических фраз может и не быть правильным фрагментом. В связи с этим введем следующее определение.

0. 6. 10. Всякая непустая совокупность  $T$ -диагностических фраз, являющаяся правильным фрагментом, называется *диагностическим фрагментом* для типа  $T$ .

В частности,  $T$ -диагностическим фрагментом в  $S$ - $T$ -правильном языке является совокупность всех  $T$ -диагностических фраз. Этот фрагмент мы назовем *главным  $T$ -диагностическим фрагментом* типа  $T$ . В дальнейшем мы будем предполагать, что тип  $T$  фиксирован (в интерпретации это будет один из типов существительных), поэтому мы будем говорить не о  $T$ -диагностических контекстах, а просто о диагностических контекстах и т. п., опуская  $T$  в последующих определениях.

Теперь мы хотим использовать главный (для любого из типов существительных) диагностический фрагмент с целью определения падежа, т. е. определить падеж через понятие класса равноемких (относительно этого фрагмента) контекстов.

Каков лингвистический смысл исключения всех остальных контекстов? Контексты, которые допускают лишь некоторые парадигмы существительного — это те контексты, которые содержат слова, согласуемые с существительными по роду.

Введя диагностические фразы, мы избавляемся в них от влияния рода и поэтому в них более четко проявляется категория падежа.

Итак, мы хотим теперь в качестве падежеобразующих взять только диагностические контексты.

Посмотрим, можно ли при этом сохранить процедуру разбиения множества диагностических контекстов на классы равноемких контекстов. На первый взгляд мы получим как раз то распределение по падежам, которое мы разбирали в начале параграфа, ибо выбранные там контексты — диагностические. Однако и при этом останутся некоторые неудобства.

Падежей все же окажется слишком много. И вот почему.

Во-первых, останутся три разных класса падежеобразующих контекстов там, где традиционная лингвистика с достаточным основанием выделяет всего один именительный падеж, ср. контексты (*вот, А*), (*А, идет*) и (*А, идут*).

Во-вторых, остаются указанные выше трудности с такими контекстами, как (*несу, А*), (*даю, А*), которые остаются диагностическими для существительных по данному нами определению. Правда, можно было бы потребовать, чтобы в число диагностических фраз не включались такие явно эллиптические, как *несу рукой* или *несу собаке*. Однако та же проблема возникает и в связи с такими возникающими из диагностических фраз контекстами, как:

(*он был, А*) — допускает *поэтом* и *поэт*

(*А, относит лодку*) — допускает *течение, течением*

(*этот человек полон, А*) — допускает *мыслей* и *мыслями*

(*он швырял в нее, А*) — допускает *камнями* и *камни*

Наличие подобных контекстов не является особенностью русского языка. Из чешского языка можно привести такие контексты, как:

(*učiti se, А*) допускает *песо* и *песему*

Из болгарского языка можно привести (*бог да, прости*) допускает *ги* и *им*.

Из польского языка можно привести (*chcieć, А*) допускает *соś* и *czegoś*, или (*blogostawić, А*) допускает *komuś* и *kogoś*.

Поэтому мы будем исходить из диагностических контекстов (напомним, что мы берем контексты, диагностические для типов существительных), но лишь некоторые из них считать падежеобразующими контекстами, т. е. контекстами, входящими в падежеобразующий класс.

**0. 6. 11.** Соберем все равноемкие диагностические контексты в классы  $K_1, K_2, \dots, K_n$  и назовем их диагностическими классами. Диагностический класс  $K$  не является падежеобразующим, если для любого слова  $x$ , допускаемого контекстами этого класса, найдется диагностический класс  $K_1$  такой, что  $K \neq K_1, K \Rightarrow K_1$  и  $K_1$  допускает  $x$ . Все остальные диагностические классы называются **п а д е ж е о б р а з у ю щ и м и**.

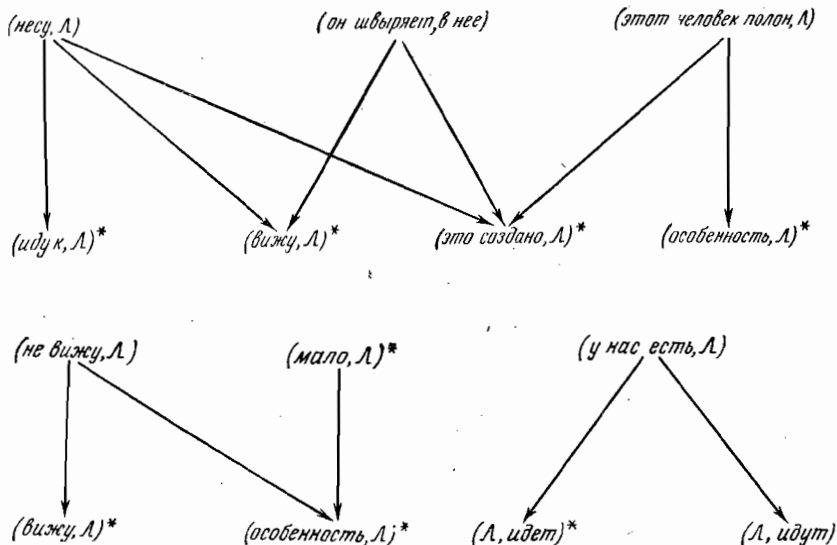
Заметим, что произведенное нами сокращение числа падежеобразующих контекстов вполне естественно: наше определение должно охватывать не все контексты, а все слова. Мы хотим, чтобы каждое слово допускалось хотя бы одним падежеобразующим контекстом. Мы выбросили такие контексты, которые являются, так сказать, «избыточными», ибо интересующие нас слова все равно допускаются другими диагностическими контекстами.



Теперь, когда дано определение падежеобразующего класса, определение 6.7 стало вполне корректным. Повторим его в более простой форме:

0. 6. 7". Падежом называется совокупность слов, допускаемая одним классом равноемких падежеобразующих контекстов.

Покажем теперь на примере, как действует наше определение для русского языка. В русском языке отношении «большей емкости» дает, в частности, следующие схемы:



Примечание: Звездочками отмечены контексты, которые по определению 6.11 являются падежеобразующими.

На первой схеме падежеобразующими по определению 6.11 являются контексты *(иду к, Л)* — соответствует дательному падежу и *(вижу, Л)* — соответствует винительному падежу, *(это создано, Л)* — соответствует творительному падежу и *(особенность, Л)* — соответствует родительному первому падежу.

Посмотрим теперь, почему контексты *(несу, Л)*, *(он швыряет в нее, Л)* и *(этот человек полон, Л)* не являются падежеобразующими. Дело в том, что они являются более емкими, чем диагностические классы, которые допускают любое из допускаемых данными контекстами слов. Мы видим, что наше определение действительно удобно.

Перейдем к другим схемам, которые мы будем комментировать более кратко. На второй схеме падежеобразующими, по определению 6.11., являются контексты *(особенность, Л)*, *(мало, Л)* и *(вижу, Л)*. Класс *(мало, Л)* является падежеобразующим потому, что не все слова, допускаемые им, допускаются менее емкими

диагностическими контекстами (например, таково слово *сахару*). Класс (*мало*,  $\Lambda$ ) соответствует родительному второму падежу.

Вторая схема показывает, что выделенный нами ранее специальный класс контекстов типа (*не вижу*,  $\Lambda$ ) не является падежеобразующим. Он более емкий, чем другие диагностические классы, например: (*вижу*,  $\Lambda$ ) и (*особенность*,  $\Lambda$ ) и любое слово, допускаемое контекстом (*не вижу*,  $\Lambda$ ), например, *стола*, допускается и одним из менее емких диагностических классов.

Третья схема дает два разных падежеобразующих класса ( $\Lambda$ , *идет*) и ( $\Lambda$ , *идут*). Дело в том, что более емкий контекстный класс (*у нас есть*,  $\Lambda$ ) не допускает других слов, кроме тех, которые допускаются одним из этих двух классов.

Это показывает, что нет принципиальной разницы между характером второй и третьей схем, хотя одна из них дает результат, согласующийся с традицией, а другая — нет. Таким образом, как бы мы ни изменяли определения, мы все равно получим результат, отклоняющийся от желаемого. Если определение строить так, чтобы падежеобразующим был в этом случае более емкий класс, то мы должны были бы выделить специальные падежи, соответствующие контекстам класса (*не вижу*,  $\Lambda$ ), или же (*несу*,  $\Lambda$ ), но зато при этом в русском языке будет лишь один именительный падеж. Вряд ли стоит на это идти, ибо именительные падежи можно объединить специальной процедурой (см. § 41).

Сделаем теперь несколько замечаний относительно принятого нами определения падежа (а именно определения 6.7').

Прежде всего очевидно, что падеж есть категория синтагматическая, ибо если два слова входят в одно семейство, то допускаются теми же контекстами (в том числе и теми же диагностическими).

Второе замечание касается следующего обстоятельства. Наше определение падежа несколько отходит от лингвистической традиции тем, что в нем категория падежа может быть установлена и для аморфных слов, а также для таких, где все формы парадигмы входят, например, в одно семейство. Лингвисты справедливо считают, что в этом случае нельзя говорить о том, что в языке есть падежи.

В связи с этим введем следующее разделение между собственными и несобственными падежами.

**0.6.12.** Категория падежа  $C_1$  называется собственной, если существуют  $x$  и  $y$  такие, что а)  $x \in \Gamma(y)$ , б)  $x \in C_1$  и  $y \notin C_1$  и в)  $y$  допускается некоторым классом падежеобразующих контекстов (т. е.  $y$  входит в какой-то другой падеж, чем  $x$ ).

Далее везде, где говорится о падеже (в том числе и об абстрактном падеже), будут иметься в виду собственные падежи.

Наконец, необходимо сделать одно замечание относительно определения падежа в некотором контексте. Дело в том, что определенные нами категории падежей, как мы уже отмечали,

могут пересекаться, т. е. одно слово может входить в два разных падежа. Мы до сих пор определяли не то, какой падеж имеет данное вхождение слова во фразу, а в какие падежи вообще входит данное слово.

Однако предложенная нами процедура легко может быть видоизменена так, чтобы определить падеж слова в контексте.

**0. 6. 13.** Слово  $x$  имеет падеж  $C$  в контексте  $(f, g)$  в следующих случаях:

а) если  $(f, g)$  входит в класс падежеобразующих для  $C$  контекстов;

б) если  $(f, g)$  является менее емким, чем контексты падежеобразующего для  $C$  класса;

в) если  $(f, g)$  является более емким, чем контексты падежеобразующего для  $C$ -класса, но слово  $x$  допускается этим классом.

Теперь мы можем определить, какой падеж имеет слово *брата* в контекстах: (*вижу большого*,  $\Lambda$ ), (*два*,  $\Lambda$ ), (*несу*,  $\Lambda$ ).

По пункту б) оно имеет винительный падеж, ибо данный контекст менее емкий, чем контекст (*вижу*,  $\Lambda$ ). Слово *брата* в контексте (*два*,  $\Lambda$ ) имеет первый или второй родительный падеж (наше определение не дает здесь однозначного решения). Слово *брата* в контексте (*несу*,  $\Lambda$ ) имеет по пункту в) винительный падеж, ибо оно допускается контекстом (*вижу*,  $\Lambda$ ) и этот контекст менее емкий, чем контекст (*несу*,  $\Lambda$ ).

#### § 41. СУБКТЕГОРИИ.

##### ВОПРОС О КАТЕГОРИЯХ ЧИСЛА СУЩЕСТВИТЕЛЬНОГО

Мы видели, что в один падеж могут входить несколько слов, относящихся к той же парадигме, например, *стола*, *столов* или *женой*, *женою*, *женами*. Такие объединения различны не только по существу, но и формально, в связи с чем появляется надежда разграничить эти случаи при помощи формального определения.

Вообще говоря, для данных пар достаточно очень простого разграничения, а именно достаточно заметить, что *женой*  $\in S$  (*женою*), в то время как *стола*  $\notin S$  (*столов*).

Мы примем, однако, более сложную систему определений, которая позволит нам с единой точки зрения трактовать существительные и глаголы и полный лингвистический смысл которой прояснится лишь в § 44.

**0. 6. 14.** Назовем дополнение любого  $T_1$ -диагностического фрагмента  $\delta$  до всего множества правильных фраз (т. е. совокупность  $\bar{\delta}$  всех тех правильных фраз, которые не входят в диагностический фрагмент  $\delta$ ) —  $T_1$ -антидиагностическим фрагментом.

Лингвистический смысл этого понятия в том, что, определив падеж при помощи фраз фрагмента  $\delta$ , мы обращаемся к остаю-

щейся совокупности фраз и пытаемся зафиксировать те различия, которые не связаны с падежом. Для того, чтобы избавиться от влияния рода, мы будем рассматривать при этом лишь формы одной парадигмы. Все эти соображения (а также структура материала, разбираемого далее в § 44) приводят к следующему определению.

0. 6. 15. Два слова  $x$  и  $y$ , входящие в один падеж и такие, что  $x \in \Gamma(y)$  и  $\Gamma(x) \subseteq T_1$ , принадлежат к одной субкатегории, если существует  $T_1$ -антидиагностический фрагмент  $\delta$ , на котором имеет место  $x \xrightarrow{\delta} y$  или  $y \xrightarrow{\delta} x$ .

Применяя данное определение, мы сразу устанавливаем, что *женой* и *женою* относятся к одной субкатегории, ибо *женой*  $\in S$  (*женою*), фрагмент, не содержащий контекстов, равноемких с (*доволен*,  $\Lambda$ ), есть антидиагностический, причем на нем *женою* взаимозаменяемо с *женой*.

С другой стороны, *женой* и *женами*, *стола* и *столов* попарно относятся к разным субкатегориям, ибо любой антидиагностический фрагмент будет содержать фразы вроде *доволен своей женой*, *доволен своими женами*, *два больших стола*, *пять больших столов*.

Итак, мы как будто получили процедуру для выделения форм единственного и множественного числа, ибо понятие парадигмы практически не влияет на определение падежа. Это в известной мере противоречит принципу, выдвинутому в § 15 и гласящему, что категории числа должны задаваться в славянских языках извне.

На самом деле здесь никакого противоречия нет. Даже если считать, что объединение в одной парадигме форм единственного и множественного числа может целиком диктоваться внутриязыковыми соображениями (ср. рассуждения, приводимые в § 35), мы должны констатировать, что при помощи понятия субкатегории мы получили возможность относить к одному числу или к разным числам лишь словоформы в пределах одного падежа. У нас нет никаких средств для выяснения того, принадлежат ли к одному или к разным числам слова *стола* и *столом* или *стола* и *столами*. Таким образом, извне обязательно должна быть задана информация о том, как объединять субкатегории, полученные в пределах разных падежей.

Категория числа строится как объединение некоторых субкатегорий (на основании внешнего критерия, нами не рассматриваемого).

Вернемся теперь к рассматривавшейся в предыдущем параграфе ситуации, когда выделялось два разных падежеобразующих класса: ( $\Lambda$ , *идет*) и ( $\Lambda$ , *идут*) и, стало быть, два разных падежа.

Легко видеть, что в таком случае система категорий была бы лишена свойства категориальной однородности (см. § 39).

В самом деле, сравним следующие пучки категорий:

(МО, ед., первый именит. падеж)	<i>мальчик</i>
(МО, ед., родительный падеж)	<i>мальчика</i>
(МО, мн., родительный падеж)	<i>мальчиков</i>
(МО, мн., первый именительный падеж)	отсутствует

Если же объединить первый именительный и второй именительный в одном падеже, то система будет категориально однородной.

Поэтому мы можем считать, что на определенном этапе происходит такое минимальное укрупнение категорий падежа, при котором язык становится категориально однородным.

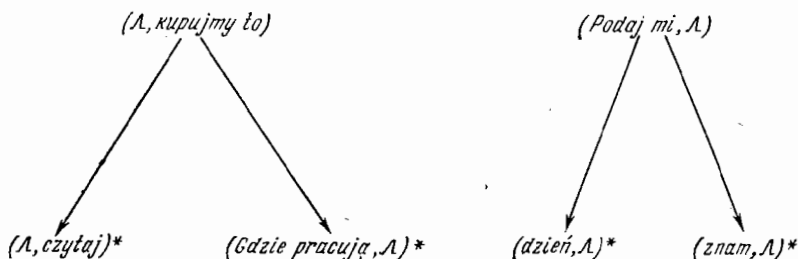
### § 42. ПАДЕЖИ В ДРУГИХ СЛАВЯНСКИХ ЯЗЫКАХ

В польском языке можно выделить следующие классы диагностических контекстов

Класс контекстов	Допускаемое слово
1. (Gdzie pracuje, $\Lambda$ )	pan, chłop, stół, matka, woda, jezioro
2. (Gdzie pracują, $\Lambda$ ), ( $\Lambda$ , czytajcel)	panowie, chłopci, stoły, matki, wody, jeziora
3. (Byłem bez, $\Lambda$ ), (dzień, $\Lambda$ )	pana, chłopca, stoła, matki, wody, panów, matek, wod, jezior
4. (Podaj mi, $\Lambda$ )	pana, chłopca, stół, matku, wody, wodu, jeziora, panów, chłopów. stołów, matek, wod, jezior
5. (Oddaj list, $\Lambda$ )	panu, chłopu, stołowi, matke, wode, jeziору, panom, chłopom, stołom, matkom, wodom, jeziorom
6. (znam, $\Lambda$ )	pana, chłopca, stół, matkę, wodę, jezioro panów, chłopów, stołów, matki, wody, jeziora
7. (Byłem w, $\Lambda$ )	panu, chłopce, stołe, wode, matke, jeziore, panach, chłopach, stołach, matkach, wodach, jeziorach
8. (Mówiłem s, $\Lambda$ )	panem, chłopem, stołem, matką, wodą, jeziorem, panami, chłopami, stołami, matkami, wodami, jeziorami
9. ( $\Lambda$ , czytaj)	pane, chłopce, stołe, matko, wodo, jezioro

10. (Λ, kupujmy to)

pane, chłope, stote, matko, wodo,  
jezioro, panowie, chłopi, stoły,  
matki, wody, jeziora



Здесь контексты типа 4 не являются падежеобразующими также, как и русские контексты типа (дай мне, Λ), ибо любое из слов, допускаемых ими, допускается и каким-нибудь из более емких контекстов. Интерес представляют еще словоформы типа *walca*, *poloneza*, *twista*, допускаемые контекстами 3, 4, 6, поскольку до сих пор не решен спор, имеем ли мы здесь дело с винительным или родительным падежами<sup>4</sup>. По определению 6. 13, пункту в) в контексте 4 слово *walca* имеет и винительный падеж и родительный, поскольку оно допускается обоими классами 3 и 6, менее емкими чем 4. Этот довольно неожиданный результат вполне согласуется с двойственным положением таких словоформ в указанных контекстах.

Не совсем похоже на русский язык и соотношение по емкости контекстов типа 10, 9 и 2, где более емкими оказываются «звательные» контексты.

Получается, во-первых, два именительных, но мы уже показали, что в таких случаях для достижения категориальной однородности целесообразно объединить два именительных. Во-вторых, категориальная неоднородность возникает из-за того, что имеется звательный падеж только для единственного числа. Исходя из той же необходимости категориальной однородности, можно разбить контекстный класс 2. (Gdzie pracują, Λ), (Λ, czytajcie) на два падежа, как и поступают в традиционной грамматике.

В сербском, словенском, верхнелужицком языках положение такое же, как и в польском, если отвлечься от контекстов типа 4.

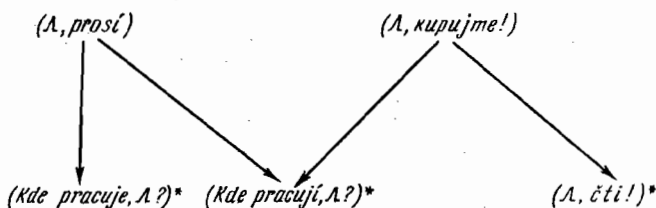
Несколько иную картину дает чешский язык в связи с омонимией третьего лица единственного и множественного числа для целого ряда глаголов.

<sup>4</sup> См.: L. Hjelmslev. *Animé et inanimé, personnel et non-personnel*. «Travaux de l'Institut de linguistique», vol. 1. Paris, 1956; см. также: A. Heinz. *System przypadkowy języka polskiego*. Kraków, 1965, стр. 95.

В чешском языке полезно рассмотреть следующие классы контекстов.

Класс контекстов	Допускаемое слово
1. (Λ, prosí)	žena, kůň, přítel, hrad, ženy, koně
2. (Kde pracuji, Λ), (Λ, spojte se!)	ženy, koně, přátelé, hrady
3. (Kde pracuje, Λ?)	žena, kůň, přítel, hrad
4. (Byl jsem bez, Λ), (den, Λ)	ženy, koně, přitele, hradu, žen, koňů (koní), přítel, hradů
5. (Napiši, dopis)	ženě, příteli, hradu, ženám, hradům
6. (Znam, Λ)	ženu, koně, přitele, hrad, ženy, hrady, přátele
7. (Mluvil jsem s, Λ)	ženou, koněm, přítelem, hradem, ženami, koňmi, přáteli
8. (Mluvil jsem o, Λ)	ženě, koni, příteli, hradě, ženách koních, přátelích
9. (Λ, čti!)	ženo, koni, příteli, hrade
10. (Λ, kupujme!)	ženo, koni, příteli, ženy, koně, přátelé

Если в чешском языке применить ту же процедуру объединения двух именительных и введения звательного множественного числа, то мы получим те же падежи, что и в польском. Таким образом, чешский язык отличается схемой большей емкости контекстов.



Представляет интерес применить выработанную систему определений к современному болгарскому языку, ибо вопрос о том, существуют ли вообще в болгарском языке падежи, подвергается оживленной дискуссии. Как известно, в болгаристике по этому вопросу существуют разные точки зрения.

Обратившись к диагностическому множеству фраз современного болгарского языка, ориентировочно получаем следующее.

- |                   |  |
|-------------------|--|
| 1. а) (Λ, гледи)  | <i>Иван, брат, командир(ът), той, тя, коза(та), куча(та), майка(та),</i> |
| б) (Λ, гледат)    | <i>братя(та), командирити(те), сучета(та), те, кози(те), майки(те)</i>   |
| 2. а) (слушай, Λ) | <i>Иване, майко, народе</i>  |
| б) (слушайте, Λ)  | <i>Ивани, майки, народи</i>  |
| 3. (покани ли, Λ) | <i>Ивана, Иван, мене, ме, него, го, тях, ги, нас, ни</i>                 |
| 4. (казах, Λ)     | <i>Ивану, мене, ми, им</i>   |

После объединения контекстных классов 1а и 1б, 2а и 2б (в случае звательного в отличие от чешского, польского и сербского контексты единственного и множественного числа различаются в связи с определенностью в им. п.) — по тем же соображениям, что и для двух именительных русского языка, получаем следующие падежи болгарского языка

1. Именительный
2. Звательный
3. Винительный
4. Дательный

Обратимся теперь к основному вопросу, волнующему грамматистов. Как известно, для большинства существительных современного болгарского языка в контексты класса 4 подставляются не отдельные слова, а сочетания существительного с предлогом. Можно ли считать эти словосочетания падежными формами?

Принятая нами концепция требует, во-первых, чтобы указанные выше контексты действительно были диагностическими, т. е. допускали любую парадигму существительного, например, контекст (*казах. Λ*) допускал какую-то форму из парадигмы *Г* (*жена*). Назвав соответствующие контексты диагностическими, мы неявно предположили, что *на жената* есть элемент из парадигмы *Г* (*жена*), и т. п. Если принять эту точку зрения, то мы с необходимостью придем далее к выводу о наличии падежей в болгарском языке.

Возможна, однако, и иная точка зрения, а именно, что такие контексты, как (*покани ли, Λ*) или (*казах. Λ*) вообще не являются диагностическими для существительных болгарского языка, поскольку они допускают не любую парадигму существительного. Они остаются диагностическими лишь для местоимений, поэтому только местоимения и имеют падежи в болгарском языке. Таким образом, вопрос о существовании падежей упирается в вопрос о диагностичности контекстов. Часто, однако, применяется совсем



другая аргументация, сводящаяся в наших терминах к следующему. Допустим, что данные контексты являются диагностическими только для местоимений-существительных. Тогда мы получим падежи местоимений. Но наряду с краткими формами местоимений в этих контекстах употребляются полные формы с предлогом *на*<sup>5</sup>. Поскольку падеж однозначно определяется классом равноемких контекстов, а контексты, допускающие слова и словосочетания с *на*, одни и те же, то следует признать и эти словосочетания падежными формами.

Но при такой аргументации имеется одна трудность, с которой необходимо считаться при типологическом анализе. Мы сказали, что форма местоимения, например *ми* допускается тем же контекстом, что и сочетание *на мене* или *на брат*. Это утверждение может быть оспорено, если учесть следующее обстоятельство.

Разграничение падежей существительного проводится здесь на основе сравнения с так называемыми «безударными» местоимениями (формы ударных местоимений обнаруживают значительно больший параллелизм с формами имени, ср. в болгарском: *на мене* — *на Иван*). К этому сводится в конце концов указанная выше аргументация Е. В. Чешко.

Между тем «безударные местоимения» в болгарском (а также в сербском) представляют собой по сути дела связанные морфы, образующие часть единой глагольной формы. Это явление, отмеченное разными авторами<sup>6</sup>, имеет своим следствием более или менее (для разных языков по разному) фиксированное положение безударного местоимения во фразе, в то время как существительное более свободно, а в связи с этим трудно говорить и об одинаковой дистрибуции имени существительного и местоимения<sup>7</sup>.

Однако в нашем анализе мы отвлечемся от этого явления, тем более, что оно еще недостаточно изучено, и будем считать, что безударное местоимение и группа «существительное с предлогом» имеют одинаковую дистрибуцию.

---

<sup>5</sup> Е. В. Чешко. Об изучении функций предлогов в болгарском языке. — КСИС, вып. 10. М., 1953; Она же. Падежи и предлоги в современном болгарском литературном языке. «Основные вопросы грамматики болгарского языка». М., 1960.

<sup>6</sup> Особенно важными представляются замечания Теньера (L. Tesnière. *Éléments de syntaxe Structurale*. Paris, 1959). См. также: Г. А. Дыхун. Синтаксис местоименных клитик в болгарском и македонском литературных языках. Автореф. канд. дисс. Л., 1966.

<sup>7</sup> Втягивание безударного местоимения в состав агглютинативной глагольной формы в болгарском языке (особенно народном) наглядно подтверждается и так называемым плеонастическим употреблением личных (безударных) местоимений, например, *Помниш ли я ти Милка 'помнишь ли ее ты Милку'* (см.: Ю. С. Маслов. Очерк болгарской грамматики. М., 1956, стр. 145).

## § 43. АБСТРАКТНЫЙ ПАДЕЖ КАК ОСНОВНАЯ НЕПАРАДИГМАТИЧЕСКАЯ КАТЕГОРИЯ

Рассмотрим теперь вопрос о том, что дает абстрактное определение падежа при переносе его на другие части речи. Ясно, что местоимения-существительные во всех славянских языках, кроме болгарского, имеют те же падежи, что и существительные, поскольку употребляются в тех же контекстах.

Попробуем применить теперь абстрактное определение падежа, данное нами, к другим частям речи. Правда, здесь возникает чисто терминологическая трудность, состоящая в том, что для некоторых частей речи термин «падеж» придется употреблять в необычном смысле. Впрочем, путаница создается здесь из-за традиционного словоупотребления. То, что обычно называют «падежом прилагательного», представляет собой с формальной точки зрения функцию совсем иной структуры, чем падеж существительного.

Для преодоления этой терминологической трудности мы будем пользоваться следующей терминологией. Категорию, определенную в 0. 6. 7'', мы будем отныне называть абстрактным падежом или казусом, а соответствующий класс контекстов будем называть казусным. Таким образом, для существительных и местоимений-существительных понятия падежа и казуса совпадают, а для других частей речи они могут различаться.

Обратимся к прилагательным. Недиагностическими будут только контексты вида (*я видел, книгу чем та*), т. е. те, которые допускают формы сравнительной степени, ибо такие формы имеются не во всех парадигмах (Впрочем, возможна точка зрения, что формы с *более* входят в ту же парадигму, т. е., например, *более красивый* есть форма слова *красивый*). Эта трактовка вполне аналогична трактовке болгарских существительных в § 42).

Остальные контексты будут диагностическими. Казусных классов будет столько, сколько имеется разных форм прилагательных и, например, к разным казусам будут относиться *маленькую, маленького и маленьких*, ибо контексты (*вижу, книгу*), (*вижу, мальчика*) и (*вижу, мальчиков*) допускают разные слова.

Ничего не дает в случае прилагательных и понятие субкатегории, ибо во фразах антидиагностического фрагмента слова вроде *маленькую* вообще не встречаются.

В русском языке казус прилагательного практически совпадает с семейством. Таково положение в большинстве славянских языков. Исключение составляет чешский язык, где, например, контекст (*А, žak prasijí*) допускает *sízi* и *novú*, но эти слова принадлежат двум разным семействам (связано с омонимичностью форм мягкого склонения).

Почему же все-таки говорят о падеже прилагательных как о чем-то схожем с падежом существительных? По-видимому, потому, что прилагательное имеет такие же падежи, как существи-

тельное в эллиптических цепочках, причем если прилагательное имело в некоторой эллиптической фразе падеж *С*, то ту же форму оно сохраняет и при употреблении с существительным, имеющим падеж *С*. Иначе говоря, слово *большую* имеет в контексте (*вижу, девочку*) винительный падеж потому, что оно имеет тот же падеж, что и существительное *девочку* в контексте (*вижу, Λ*), сохраняя эту же форму в контексте, где уже есть слово *девочку*, употребленное в винительном падеже.

Однако, если допустить, как это было сделано в § 29, эллиптические фразы, то прилагательные будут иметь, во-первых, все те Kasusы, которые соответствуют падежам существительного и, во-вторых, все Kasusы, соответствующие семействам. Это показывает, что введение в рассмотрение эллиптических фраз не решает автоматически проблемы падежа прилагательного.

Напомним теперь, что при рассмотрении и эллиптических фраз имеют место соотношения замещения вроде следующих:

*стол* ⇒ *маленький*  
*лампа* ⇒ *маленькая*  
*окно* ⇒ *маленькое*  
*стола* ⇒ *маленького*  
*окна* ⇒ *маленького*  
*лампы* ⇒ *маленькой* и т. п.

Это обстоятельство наводит на мысль следующим образом определить аналог падежа прилагательного.

0. 6. 16. Назовем *с у н е р к а з у с о м*, подчиненным некоторому Kasusу *С*, совокупность слов *x* таких, что  $y \Rightarrow x$ , где  $y \in C$ .

Тогда *маленький*, *маленькая* и *маленькое* объединяются в один суперKasus, подчиненный именительному падежу.

Переходим теперь к понятию Kasusа применительно к глаголам.

Если согласиться с принципом, выдвинутом в § 35, согласно которому формы вида попадают в разные парадигмы, то недиагностическими будут для глагола контексты типа (*буду, Λ*), (*начинаю, Λ*), (*продолжаешь, Λ*), и т. п., т. е. те, которые допускают лишь инфинитивы несовершенного вида.

Диагностическими будут контексты вида

(*я, Λ*)      (*я, книгу*)      (*я, ему*)      (*я, городом*)  
(*ты, Λ*)  
(*он, Λ*)      и т. п.      и т. п.  
(*она, Λ*)      и т. п.  
(*оно, Λ*)  
(*мы, Λ*)  
(*вы, Λ*)  
(*они, Λ*)

В этом примере контексты, выписанные в одну строку, различаются тем, что допускают слова, относящиеся к разным категориям управления — мы будем условно называть эти категории классами управления. Таким образом, классы управления суть синтагматические категории глагола.

Теперь обратимся к классам контекстов, выписанным в один столбец. В множество слов, соответствующее одной категории, попадает более одной формы одной парадигмы (например, *читаю, читал, читала*). Иначе говоря, мы имеем здесь с формальной точки зрения такое же положение, какое было в падежах существительных, где в множество, соответствующее одному падежу, попадали две формы числа. Это обстоятельство привело, как мы помним, к формированию понятия субкатегорий. Применим теперь соответствующую процедуру к глаголу.

#### § 44. ИССЛЕДОВАНИЕ КАТЕГОРИЙ ГЛАГОЛЬНОГО ЛИЦА И РОДА В СЛАВЯНСКИХ ЯЗЫКАХ

Рассмотрим одновременно два заданных извне разбиения на правильные фрагменты. В первом разбиении задаются следующие фрагменты:

$\Delta_{11}$  Совокупность всех тех правильных фраз, в которых могут употребляться в именительном падеже лишь местоимения *ты* и *вы*, т. е. совокупность высказываний, которые выполняют по Бюлеру <sup>8</sup> функцию призыва (Appel)

$\Delta_{12}$  Совокупность всех тех правильных фраз, в которых могут употребляться в именительном падеже лишь местоимения *я* и *мы*, т. е. совокупность высказываний, которые выполняют по Бюлеру функцию самовыражения (Ausdruck)

$\Delta_{13}$  Совокупность всех тех фраз, в которых не могут участвовать в именительном падеже местоимения *я*, *ты*, *мы*, *вы*, т. е. те, которые выполняют, по Бюлеру, функции дескрипции (Darstellung).

$\Delta_{14}$  Смешанные типы.

Во втором разбиении задаются следующие фрагменты:

$\Delta_{21}$  Множество фраз, в которых употребляются в именительном или звательном падеже лишь названия лиц мужского пола, т. е., например, исключаются фразы вида: *ты видишь, мама*.

$\Delta_{22}$  Множество фраз, в которых употребляются в именительном или звательном падеже лишь названия лиц женского пола, т. е., например, исключены фразы вида: *я, отрок, зажигаю свечи*.

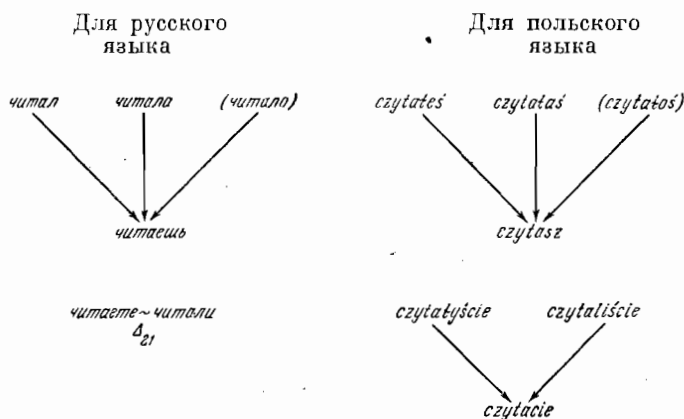
$\Delta_{23}$  Остальные фразы.

<sup>8</sup> К. В ü h l e r. Sprachtheorie. Jena, 1934.

Возьмем для примера лишь одну из парадигм, а именно  $\Gamma$  (*читать*) и для сравнения соответствующие польские формы<sup>9</sup> и будем изучать отношения замещения, возможные на соответствующих фрагментах.

Каждый из приведенных фрагментов является антидиагностическим для любого глагольного типа. Фрагмент  $\Delta_{11}$  можно представить как дополнение всего множества отмеченных фраз до фрагмента, включающего все фразы вида  $fxgyh$ , где  $x$  — местоимения *я, мы, он, она, оно, они*, а  $y$  — соответствующая глагольная форма. Разумеется, контекст  $(fxg, h)$  допускает любую глагольную парадигму данного типа. Этот фрагмент — правильный, так как он получен удалением целых семейств. Аналогичное рассуждение можно провести для всех других приведенных фрагментов.

Итак, мы имеем дело с фрагментами, для которых имеет смысл оперировать с понятием субкатегории. На фрагменте  $\Delta_{11}$  мы получим следующие схемы



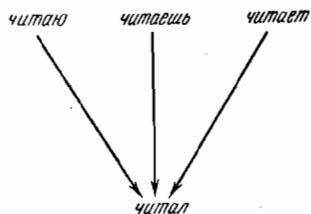
Аналогичные схемы мы получим на фрагментах  $\Delta_{12}$  и  $\Delta_{13}$ , только замещающей будет форма 1-го и соответственно 3-го лица.

Перейдем ко второму разбиению. На фрагменте  $\Delta_{21}$  имеются следующие схемы замещения

<sup>9</sup> Близкое к польскому распределение форм дает чешский язык; о синтетичности форм претерита в польском языке и отчасти в чешском см.: Fr. Корейну. Syntetičnost polského i českého préterita. «Studia z filologii polskiej i słowiańskiej», 5. Warszawa, 1965, стр. 341—346. — С другой стороны, необходимо подчеркнуть, что представление форм претерита в польском языке как синтетических является определенной идеализацией (см. § 36) из-за возможности трактовки соответствующего показателя как подвижной частицы. Но в этом случае речь может идти о расширенной парадигме в смысле § 36.

Для русского языка

Для польского языка

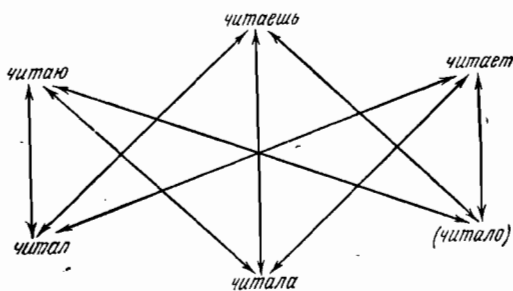


czytam ~ czytałem  
 czytaś ~ czytałeś  
 czyta ~ czytał  
 czytaliśmy ~ czytamy  
 czytaliście ~ czytacie  
 czytali ~ czytają

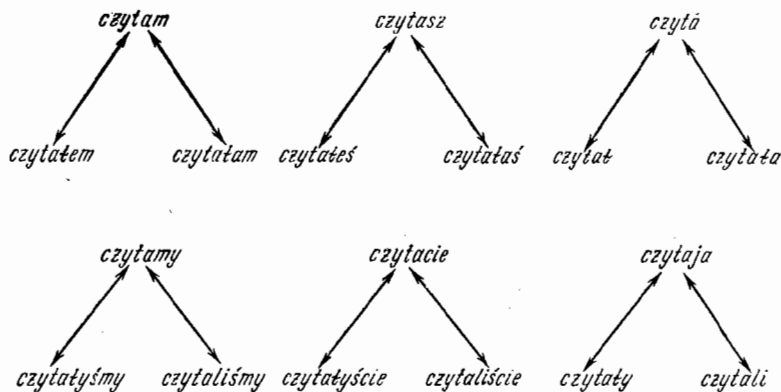
читают ~ читали  
 $\Delta_{21}$

Аналогичные схемы мы получим на фрагменте  $\Delta_{22}$  (мы их здесь выписывать не будем). Общая схема отношений ограниченного замещения в смысле определения 3. 36 выглядит следующим образом:

Для русского языка



Для польского языка



Подобные схемы наглядно показывают типологическое расхождение столь близких языков, как русский и польский. Мы привели, однако, эти схемы для того, чтобы выделить соответствующие субкатегории. Получаем следующие субкатегории:

1. *Читаю, читал, читала* соответствует 1-у л. ед. ч.
2. *Читаешь, читал, читала* соответствует 2-у л. ед. ч.
3. *Читает, читал, читала* соответствует 3-у л. ед. ч.
4. *Читаем, читали* соответствует 1-у л. мн. ч.
5. *Читаете, читали* соответствует 2-у л. мн. ч.
6. *Читают, читали* соответствует 3-у л. мн. ч.
7. *Читал, читаю, читаешь, читает* соответствует м. р. глагола.
8. *Читала, читаю, читаешь, читает* соответствует ж. р. глагола.

В польском выделяются аналогичные категории с дополнительным введением лично-мужского глагольного рода во множественном числе; это особенно интересно, если учесть, что схемы ограниченного замещения в обоих языках, как мы видели, весьма различны.

#### § 45. МОДЕЛИ ФУНКЦИЙ РОДА

В «Моделях языка» предложено моделировать категории грамматического рода через понятие подкласса. В несколько уточненной форме соответствующие понятия определяются следующим образом.

0. 6. 17. Два слова  $x$  и  $y$ , принадлежащие к одному семейству, входят в одно подсемейство (подсемейство слова  $x$  обозначим через  $PS(x)$ ), если в семействе любого слова  $z$ , входящего в ту же парадигму, что и  $x$ , найдется слово  $w$ , входящее в ту же парадигму, что и  $y$  и 2) в семействе любого слова  $u$ , входящего в ту же парадигму, что и  $y$ , найдется слово  $v$ , входящее в ту же парадигму, что и  $x$ .

Так, в русском языке слово *дом* входит в то же подсемейство, что и слово *стол*, ибо они, во-первых, входят в то же семейство и, во-вторых, в семействе любого слова из парадигмы *дом* (например, возьмем семейство слова *дому*) найдется слово, входящее в ту же парадигму, что и *стол* (в нашем примере таким словом будет слово *столу*). Слова *столу* и *судье* принадлежат одному семейству, но разным подсемействам, ибо в парадигме слова *столу* имеется слово *стол*, для которого нет  $E$ -эквивалентного в парадигме слова *судье*.

0. 6. 18. Подклассом слова  $y$ , который мы обозначим через  $M(y)$  называется совокупность всех слов  $x$  таких, что  $PS(x) \cap \Gamma(y) \neq \emptyset$ .

В «Моделях языка» доказываются следующие утверждения:

Т. 6. 4. Подклассы образуют разбиение.

Т. 6. 5. Производное от разбиения на подклассы совпадает с разбиением на типы.

Следует признать, что пользоваться данным в «Моделях языка» определением нелегко. Поэтому мы дадим здесь более простое и прозрачное определение подкласса:

0. 6. 18'. Два слова  $x$  и  $y$  относятся к одному подклассу, если а) для любого слова  $x' \in \Gamma(x)$  найдется  $y' \in \Gamma(y)$  такое, что  $x' \in S(y')$  и б) для любого слова  $y'' \in \Gamma(y)$  найдется такое слово  $x'' \in \Gamma(x)$ , что  $y'' \in S(x'')$ .

Т. 6. 6. Определения 6. 18 и 6. 18' равносильны.

Определение 6. 18' удобно и с той точки зрения, что его легко сравнить с определением грамматического рода, предложенным Маркусом<sup>10</sup>.

0. 6. 19. Два слова  $x$  и  $y$  относятся к одному роду, в смысле Маркуса, если любое слово  $x' \in \Gamma(x)$  можно соединить  $\Gamma$ - $S$ -цепью, насыщающей не более трех слов, с любым словом  $y' \in \Gamma(y)$ .

Легко показать следующее:

Т. 6. 7. Для того, чтобы два слова  $x$  и  $y$  принадлежали к одному роду, в смысле Маркуса, необходимо и достаточно, чтобы имело место хотя бы одно из двух: а) для любого слова  $x' \in \Gamma(x)$  нашлось такое слово  $y' \in \Gamma(y)$ , что  $x' \in S(y')$  или б) для любого слова  $y' \in \Gamma(y)$  нашлось слово  $x'$ , такое, что  $x' \in \Gamma(x)$  и  $x' \in S(y')$ .

Таким образом, по сути дела определение Маркуса отличается от определения 6. 18' тем, что в нем достаточно выполнения хотя бы одного из требований а) или б) из определения 6. 18'. Поэтому, в частности, роды, в смысле Маркуса, не образуют, вообще говоря, разбиения. Из теоремы 6. 7 сразу же выводится и такое следствие, что два слова, входящие в один подкласс, входят в один род в смысле Маркуса.

Моделирование рода через подкласс достаточно удобно. Оно позволяет выделить в русском языке по крайней мере шесть родов:

1. такие слова, как *герой, конь, председатель...*
2. » » » *стол, стул, край, род...*
3. » » » *женщина, мать, героиня...*
4. » » » *лампа, страна, окрестность...*
5. » » » *млекопитающее, насекомое...*
6. » » » *окно, семейство...*

Указанные роды выделяются и по модели Маркуса.

---

<sup>10</sup> См.: С. М а р к у с. Грамматический род и его логическая модель. «Математическая лингвистика» (сб. переводов). М., 1964, стр. 122—144



Разница между обоими моделями проявляется при рассмотрении *singularia pluralia tantum*. По определению 6.18, выделится еще два подкласса:

7. такие слова, как *сани, штаны, гусли, именины, каникулы, джунгли* . . .

8. такие слова, как *молоко, добро, картофель, доказательность* . . .

По модели Маркуса (и при условии, что мы ограничиваемся взаимозамещаемостью на фрагменте, не содержащем *один из которых* и т. п.) слово *именины* входит в любой из шести выделенных родов (так, например,  $\Gamma$ - $S$ -цель, связывающая *героя* и *именинами* имеет вид:  $\text{героя} \xrightarrow{\Gamma} \text{героями} \xrightarrow{S} \text{именинами}$ ).

Мы будем пользоваться понятием подкласса как более удобным.

Из определения подкласса сразу же следует, что подкласс есть категория парадигматическая, ибо два слова из одной парадигмы тривиальным образом принадлежат одному подклассу. Отсюда и род, моделируемый через подкласс, есть парадигматическая категория. Вообще говоря, подкласс не синтагматическая категория. Интересно, однако, что имеет место следующее утверждение.

**Т. 6. 8.** Подкласс есть синтагматическая категория тогда и только тогда, когда язык формально однороден.

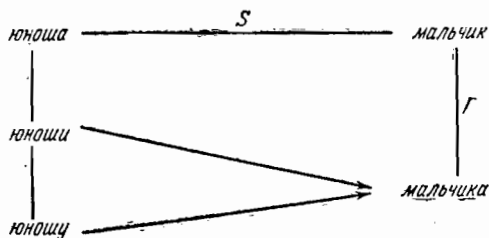
Моделирование рода через понятие подкласса имеет однако один недостаток, состоящий в том, что эта модель основана на понятии взаимозамещаемости (в основе лежит понятие семейства). Это несколько идеализирует ситуацию в естественных языках и в некоторых случаях приводит к лишним родам.

Рассмотрим, например, следующий фрагмент русского языка:

фразы: *юноша читает, мальчик читает, вижу юношу, вижу мальчика, голова юноши, голова мальчика;*

парадигмы:  $\Gamma_1 = \{\text{юноша, юноши, юношу}\}$ ,  $\Gamma_2 = \{\text{мальчик, мальчика}\}$ ,  $\Gamma_3 = \{\text{читает}\}$ ,  $\Gamma_4 = \{\text{вижу}\}$ .

Это дает следующие отношения замещения.



Здесь не только нарушается формальная однородность (ср. аналогичный пример в § 34), но по любому из приведенных опре-

делений рода *юноша* и *мальчик* относятся к разным родам. В самом деле  $\Gamma$ - $S$ -цепь между *юноши* и *мальчика* насчитывает четыре слова, слово *юноши* не эквивалентно слову *мальчика* и слова *мальчик* и *юноша* входят в разные подсемейства, а стало быть, и в разные подклассы.

Между тем естественно дать определение, которое не давало бы столь дробное разбиение на роды, не имеющее лингвистического смысла (можно добавить, что новые роды будут появляться с каждым новым случаем нейтрализации).

Поэтому мы построим еще один аналог рода.

0. 6. 20. Два слова  $x$  и  $y$  относятся к одной породе, если а) для любого  $x' \in \Gamma(x)$  найдется такое слово  $y' \in \Gamma(y)$ , что  $x' \in D(y')$  и б) для любого  $y' \in \Gamma(y)$  найдется такое слово  $x' \in \Gamma(x)$ , что  $y' \in D(x')$ .

Вопрос о том, как удобнее моделировать род, нельзя, однако, считать решенным.

Сразу же видно, что определение 6. 20 отличается от определения 6. 18 тем, что мы заменили семейства на  $D$ -классы. Это позволяет объединить в один род такие слова, как *мужчина* и *мальчик*.

В некоторых случаях однако определение 6. 20 снимает родовые различия, являющиеся существенными.

Так, по определению 6. 20 слова *мальчик* и *стол* также относятся к одному роду, ибо имеет место: *мальчик*  $\Rightarrow$  *стол*, *стола*  $\Rightarrow$  *мальчика*, а остальные формы попарно входят в одинаковые семейства.

К одному роду относятся *лампа* и *жена*, ибо имеет место *жены*  $\Rightarrow$  *лампы* и *ламп*  $\Rightarrow$  *жен*, и опять остальные формы попарно входят в одинаковые семейства. Таким образом, предложенная модель маскирует противопоставления по одушевленности—неодушевленности, которые вскрывались моделями, использующими понятие семейства, и которые, как мы увидим в дальнейшем, чрезвычайно важны для типологии славянских языков.

Поэтому мы рассмотрим еще одну модель рода. Эта модель отличается от предыдущих тем, что в ней для определения рода используется понятие падежа. В несколько других терминах она впервые была изложена А. А. Зализняком<sup>11</sup>.

А именно А. А. Зализняк предлагает исходить из того, что каждое существительное выступает в парадигме как ряд форм (с фиксированным числом и падежом).

0\*. 6. 21. (А. А. Зализняк). С о г л а с о в а т е л ь н ы м к л а с с о м называется такая совокупность существительных, что любые два ее члена, будучи взяты в любой грамматической форме (по одной и той же для обоих), «требуют» при атрибутивной связи одной и той же словоформы.

<sup>11</sup> А. А. З а л и з н я к. Категории рода и одушевленности в русском языке. — ВЯ, 1964, № 4, стр. 25—40.

Можно установить, что если бы не было омонимии форм, то совокупность слов, которые для данной фиксированной формы «требуют» при атрибутивной связи одной и той же словоформы, практически совпадала бы с семейством. В этом случае согласовательный класс Зализняка практически совпадал бы с нашим подклассом.

Наличие омонимии приводит к тому, что слова, имеющие одну и ту же форму (В, ед. ч.), например, *мальчика* и *юношу*, входят в разные семейства, но имеют одинаковое согласование. Поэтому мы попробуем, используя понятие «падежа», определенное нами ранее, так перегруппировать семейства, чтобы, расцелив омонимичные семейства, получить объединения, характеризуемые одинаковым согласованием. Введем для этого следующее определение.

**0. 6. 22.** Два слова  $x$  и  $y$  принадлежат к одному г и п е р с е м е й с т в у относительно падежеобразующего класса  $K$ , если любой контекстный класс менее емкий, чем  $K$ , допускает  $x$ , тогда и только тогда, когда он допускает  $y$ .

Возьмем вновь слова *мальчика* и *юношу*. Относительно падежеобразующего класса, содержащего контексты вроде (*вижу*,  $\Lambda$ ), эти слова входят в одно гиперсемейство, так как из менее емких контекстов такие контексты, как (*вижу маленького*,  $\Lambda$ ), допускают оба слова, а такие контексты, как (*вижу маленькую*,  $\Lambda$ ), не допускают ни того, ни другого. Таким образом, нам действительно удалось отразить идею согласования в данной фиксированной форме. Поэтому мы считаем следующее определение полным аналогом определения 6. 21.

**0. 6. 23.** Два слова  $x$  и  $y$  относятся к одному роду, в смысле Зализняка, если а) для любого падежа  $C$  и любого слова  $x' \in \Gamma(x) \cap C$  найдется слово  $y' \in \Gamma(y)$  такое, что  $y'$  принадлежит к тому же гиперсемейству относительно  $C$ , что и  $x'$ , и б) для любого падежа  $C$  и любого слова  $y \in \Gamma(x) \cap C$  найдется слово  $x' \in \Gamma(x)$  такое, что  $x'$  принадлежит к тому же гиперсемейству относительно  $C$ , что и  $y$ .

В этой форме видна явная аналогия между определением рода в смысле Зализняка и определением рода как подкласса.

Поскольку два слова, входящие в одно семейство, всегда входят в одно гиперсемейство, то два слова, принадлежащие к одному подклассу, будут входить и в один род, по Зализняку.

Для существительных определение рода, по Зализняку, дает наиболее интересные результаты, и мы будем пользоваться ими при сравнении родовых систем славянских языков. В остальных случаях, т. е. для других частей речи, целесообразнее понимать абстрактный род в смысле определения 6. 18' (не используя понятия падежа, которое недостаточно выяснено применительно к этим частям речи). Так мы и будем поступать, говоря об абстрактном роде, или генусе.

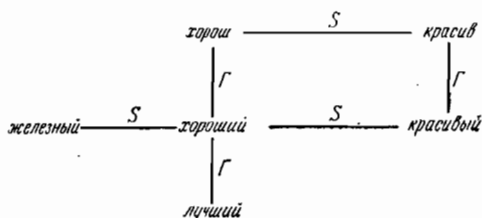
## § 46. АБСТРАКТНЫЙ РОД

### КАК ОСНОВНАЯ ПАРАДИГМАТИЧЕСКАЯ КАТЕГОРИЯ ЯЗЫКА

Хотя категории рода, как явствует из названия, моделируются как категории существительных и, действительно, как мы увидим из следующего параграфа, особенно плодотворно интерпретируются именно для этой части речи, абстрактный характер модели позволяет использовать ее для анализа парадигматических категорий, приписываемых другим частям речи.

Возьмем сначала прилагательные. Здесь, во-первых, выделяются такие слова, как *железный*, не имеющие кратких форм и, во-вторых, такие слова, как *красивый*, в парадигму которого включены формы *красив*, *красива*, *красиво*.

С другой стороны, мы имеем уже упомянутые формы вроде *хороший*, включающие в свою парадигму такие слова, как *лучший*. Получаем:



По определению 6.18' получаем три абстрактных рода, соответствующие относительным прилагательным и качественным, также разбиваемся на две группы. Более трудным для лингвистической интерпретации представляется именно это второе разделение. И все же чисто лингвистические основания для такого разделения существуют. Как тонко отметил А. В. Исаченко<sup>12</sup>, прилагательные, входящие в одну группу с *хороший*, относятся к таким, которые обозначают не объективное качество, а как бы координатную сетку, набрасываемую человеком на мир. Правда, к этому семантическому разряду относятся и некоторые другие (например, *глубокий* — *мелкий*), но из общей грамматики известно, что грамматикализации (категоризации) как правило подвергаются не все средства выражения определенной семантической категории. Возможно, что грамматическое обособление лишь некоторых противопоставлений (отражающих к тому же важнейшие оценки) имеет под собой и психологическую реальность<sup>13</sup>.

Во всяком случае особое место русского языка, формально выделяющего соответствующую группу слов, имеет типологи-

<sup>12</sup> А. В. Исаченко. Грамматический строй русского языка в сопоставлении со словацким, I. Братислава, 1954, стр. 233—234.

<sup>13</sup> Ср. опыты Осгуда и др. (G. E. Osgood, I. Suci, P. H. Tannenbaum. The measurement of meaning. Urbana, 1957), в которых выяснилось, что все пары противопоставлений, по которым испытываемые должны

ческую ценность, по-видимому, не меньшую, чем выделение, например, лично-мужского рода в существительных польского языка (рода, для осмысления которого столь же трудно найти веские основания).

Пусть теперь по определению 6.18' получены абстрактные роды существительных (совпадающие, как мы видели, с обычными грамматическими родами) и абстрактные роды прилагательных. Мы хотим теперь ввести определения, позволяющие приблизиться к традиционным родам прилагательных. Для этого, как и в § 43, мы будем рассматривать и эллиптические фразы типа *я вижу зеленую*. Введем следующие определения.

0.6.24. Мы будем говорить, что парадигма  $\Gamma(x)$  подчинена парадигме  $\Gamma(y)$ , если для любого  $x' \in \Gamma(x)$  найдется  $y' \in \Gamma(y)$  такое, что  $x' \Rightarrow y'$ , но обратное неверно (т. е. не для любого  $y' \in \Gamma(y)$  найдется  $x' \in \Gamma(x)$  такое, что  $y' \Rightarrow x'$ ).

Например, парадигма  $\Gamma(\text{большой})$  подчинена парадигмам  $\Gamma(\text{дом})$ ,  $\Gamma(\text{мальчик})$ ,  $\Gamma(\text{окно})$ ,  $\Gamma(\text{лицо})$ ,  $\Gamma(\text{женщина})$ ,  $\Gamma(\text{лампа})$ .

Введем теперь понятие «подрода, подчиненного некоторому роду»:

0.6.25. Будем говорить, что два слова  $x$  и  $y$ , входящие в один абстрактный род, входят в один п о д р о д, п о д ч и н е н н ы й р о д у  $P$ , если  $\Gamma(x)$  и  $\Gamma(y)$  не входят в  $P$  и существуют парадигмы  $\Gamma(z)$  и  $\Gamma(w)$ , входящие в  $P$  и обладающие тем свойством, что  $\Gamma(x)$  подчинена  $\Gamma(z)$ , а  $\Gamma(y)$  подчинена  $\Gamma(w)$ .

Из этого определения вовсе не следует, что подроды образуют разбиение, и это действительно не имеет места: так слово *большого* относится к тому же подроду, что и слово *большой*, и к тому же подроду, что и слово *большое*. Не следует из определения и того, что подрод подчинен ровно одному роду, и фактически одна и та же совокупность слов подчинена роду, содержащему слово *стол*, и одновременно подчинена роду, содержащему слово *мужчина*.

Итак мы видим, что определение подрода действительно отвечает особенностям функционирования тех категорий, которые обычно именуется родами прилагательного.

В применении к местоимениям-существительным определение абстрактного рода дает мало интересного. В самом деле, если иметь то разбиение на парадигмы, которое дано в § 26, то почти каждая из этих парадигм составляет и особый абстрактный род.

Определение подрода, подчиненного некоторому роду, представляет интерес лишь для парадигмы  $\Gamma(\text{он})$ .

Как мы уже указывали, каждая из таких парадигм составляет отдельный абстрактный род. Кроме того, мы знаем, что для

---

были оценить различные предметы, в сущности сводятся к трем независимым противопоставлениям, в том числе по фактору оценки (*хороший* — *плохой*). К сожалению, на славянском материале подобные эксперименты, сколько известно автору, не проводились (ср. § 62).

каждого слова *у* из парадигмы *Г* (*он*) найдется существительное, подчиняющее его. В результате этот генус, или абстрактный род, разбивается на три подрода, подчиненных соответственно мужскому (одушевленному и неодушевленному), женскому (одушевленному и неодушевленному) и среднему (одушевленному и неодушевленному). Это вполне соответствует как истории этой категории слов, так и приписыванию им форм рода в грамматике.

Обратимся к глаголу. Схема, приведенная в связи с неоднородностью, показывает, что выделяется по крайней мере два генуса, или абстрактных рода, в один из которых объединяются такие слова, как *видеть*, *вижу*; *видишь*, а в другой — такие слова, как *увидеть*, *увидю*, *увидишь* и т. п. Итак, виды глагола могут выделяться той же процедурой, что и роды существительного. Необходимо отметить, что понятие подрода, подчиненного роду, неприменимо к глаголу, так как парадигмы глагола не подчинены парадигме имени.

Между «родом прилагательного» и «родом глагола» (как они выделяются в традиционной теории) вскрывается та существенная разница, что род прилагательного в случаях субстантивации может стать подлинным родом, в то время как родовая форма глагола никогда не превращается в существительное.

Род глагола в наших терминах отражается как частный случай категории казуса глагола — эта категория с содержательной точки зрения является пересечением ряда категорий (рода, числа, лица). Но пересечением нескольких категорий (с содержательной точки зрения) оказывается и категория генуса глагола, которая различает не только виды. Пример, приведенный в § 43 в связи с рассмотрением категории казуса глагола, показывает, что синтагматическая категория, отражающая, в частности, и «классы управления», дает полную симметрию во всех формах парадигмы (*вижу* так относится к *любуюсь*, как *видишь* к *любуюсь* и т. п.). Отсюда следует, что два слова, принадлежащие к разным «классам управления», относятся к разным генусам глагола.

Таким образом, с формальной точки зрения все парадигматические категории полнозначных частей речи устроены более или менее одинаково с родом существительного.

Интересно, однако, что если парадигматические категории других гипертипов во всех славянских языках устроены более или менее одинаково (небольшие отличия русского языка в области прилагательных, отмеченные нами, достаточно спорны), то категории рода существительных в славянских языках отличаются очень большим разнообразием (см. § 47).

В заключение этого параграфа отметим следующее. В § 22 была сформулирована гипотеза, согласно которой на всем множестве фраз категориальный пучок совпадает с семейством. Теперь интересно выяснить условия, при которых имеет место такое

совпадение. Мы будем вновь исследовать этот вопрос применительно к идеализированной модели, в которой нет грамматической омонимии. Кроме того, мы еще потребуем, чтобы все категории каждого категориального пучка за исключением, может быть, абстрактной категории рода, формализуемой как категория подкласса, были синтагматическими (это условие, как мы видели, на практике всегда оправдывается).

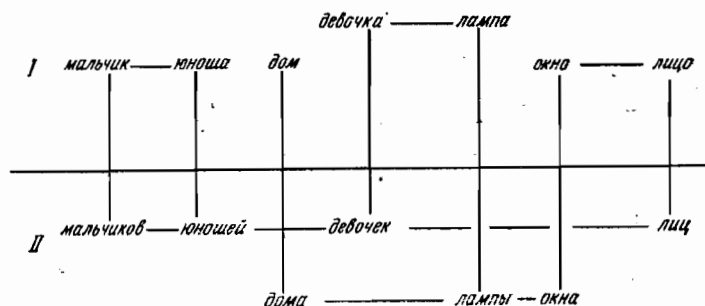
При этих условиях имеет место следующее утверждение.

**Т. 6. 9.** Если все категории, не совпадающие с подклассом, синтагматические и если язык формально однороден, то в нем два слова, относящиеся к одному семейству, относятся к одному категориальному пучку.

Мы видели в § 22, что при порождении — если принять, что правила воплощения состоят в замене категориального пучка словом — имеет место и обратное: два слова, относящиеся к одному категориальному пучку, входят в одно семейство. Поэтому если грамматика порождает однородный язык, то в нем семейство совпадает с категориальным пучком.

#### § 47. ИНВЕНТАРЬ КАТЕГОРИЙ РОДА В СЛАВЯНСКИХ ЯЗЫКАХ <sup>14</sup>

Мы условились строить типологию рода существительных, исходя из определения б. 23 (род в смысле Зализняка). Оказалось, что для выделения родов в славянских языках достаточно рассмотреть лишь некоторые гиперсемейства относительно именительного падежа (на схемах, приводимых ниже, им будет соответствовать цифра I) и относительно винительного падежа (им будет соответствовать цифра II).



<sup>14</sup> Этот параграф написан на основе исследований категорий рода, выполненных О. Г. Ревзиной (Карпинской) и отчасти опубликованных ею: О. Г. Карпинская. Типология рода в славянских языках. — ВЯ. № 6, 1964; Она же. Методы типологического описания славянских родовых систем. «Лингвистические исследования по общей и славянской типологии». М., 1966, стр. 82—100.

Этой схемы достаточно, чтобы выделить в русском языке шесть родов:

МО — мужской одушевленный: *мальчик, мальчиков, юноша* . . .

МНО — мужской неодушевленный: *дом, дома* . . .

ЖО — женский одушевленный: *девочка, девочек* . . .

ЖНО — женский неодушевленный: *лампа, лампы* . . .

СО — средний одушевленный: *лицо, лиц* . . .

СНО — средний неодушевленный: *окно, окна* . . .

Такое разбиение на роды получается лишь в том случае, если в одной парадигме объединены формы единственного и множественного числа. Если же считать, что *Г* (*мальчик*) и *Г* (*мальчиков*) и т. п. — разные парадигмы, то мы получим в единственном числе четыре рода:

МО — *мальчик, юноша*;

МНО — *дом*;

Ж — женский: *девочка, лампа*;

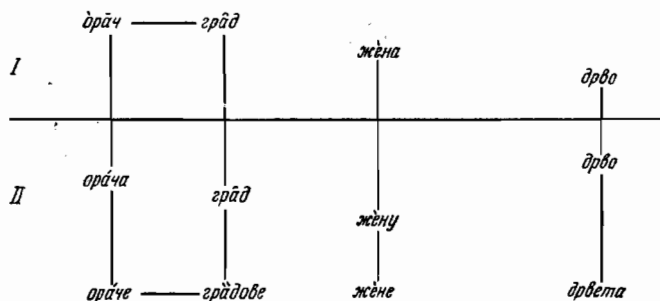
С — средний: *окно, лицо*;

и два рода во множественном числе:

О — одушевленный: *мальчик, мальчиков, девочки, девочек, лица, лиц*;

НО — неодушевленный: *дома, лампы, окна*.

Заметим, что именно такое разбиение на роды получалось в концепции, в которой единственное и множественное числа считались разными системами словоформ<sup>15</sup>.



Мы видели, что *pluralia tantum* образуют особый подкласс. Легко проверить, что *pluralia tantum* образуют и особый род, по Зализняку, который предложил называть соответствующий род «парным»<sup>16</sup>. Хотя место соответствующих слов в разных

<sup>15</sup> См.: N. Durnovo. La catégorie du genre en russe moderne. «Revue des études slaves», t. 4, № 1—2. Paris, 1924.

<sup>16</sup> А. А. Зализняк. Указ. соч.



славянских языках различно (в связи с характером разбиения на роды во множественном числе при разделенной парадигме), мы ограничимся в дальнейшем лишь шестиродовой системой для русского языка и соответствующими ей системами других славянских языков.

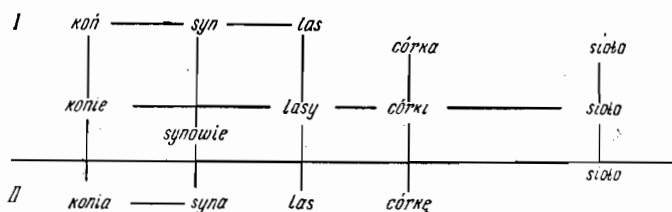
В сербском языке для выделения родов достаточно рассмотреть формы, как изображено на схеме (стр. 167).

Это дает четыре рода: МО, МНО, Ж и С.

Если мы включим в рассмотрение и приведенные на стр. 124 формы *очи* и подобные им, то в сербском языке выделяется пятый род: обоюдный, аналогичный обоюдному роду румынского языка. Поскольку, однако, таких слов очень мало, то вряд ли имеет смысл выделять в сербском языке «обоюдный род». Те же четыре рода, что и в сербском, выделяются в словенском языке, а также в чешском языке.

Разница между сербским и словенским, с одной стороны, и чешским, с другой, проявляется, если рассмотреть роды, выделяемые при разделении единственного и множественного числа на две разные парадигмы. В сербском и словенском для единственного числа выделяется тот же набор родов, что и при неразделенной парадигме, а для множественного числа получается трехродовая система (М., Ж., С.) — здесь: между прочим, наиболее четко вскрывается и разница сербского и русского языков. В чешском языке для единственного числа имеем тот же набор, что в сербском и словенском, а также в самом чешском языке при неразделенных парадигмах, для множественного числа имеем три рода (МО, С и остальные существительные, т. е. происходит склеивание родов МНО+Ж).

В польском языке имеем следующую схему:



Эта схема дает следующие роды:

ЛМ: лично-мужской *syn, syna, synowie*

НЛМО: нелично-мужской одушевленный: *koń, konia, konie*

МНО: мужской неодушевленный: *las, lasy*

Ж: женский *córka, córkę, córki*

С: средний *sioła, sioła*.

Аналогичные роды выделяются в словацком и верхнелужицком.

При отдельных парадигмах в польском языке (а также в верхнелужицком и словацком) для единственного числа выделяются роды МО, МНО, Ж и С, а для множественного числа роды ЛМ и остальные существительные.

Наконец в болгарском языке выделяется простейшая система родов: М, Ж, С, причем при разделении единственного и множественного числа в единственном числе сохраняется та же система, а во множественном родов вообще нет (все существительные принадлежат одному роду).

**Примечание.** В ряде славянских языков и диалектов намечается тенденция к стиранию противопоставления между НЛМО и МНО. Это связано с возможностью (особенно в единственном числе) параллельного употребления двух форм в винительном падеже, например в польском *17: podaj mi ten grzebień* и (в разговорной речи) *podaj mi tego grzebienia, położył papieros na stole* наряду с *położył papierosa na stole, ujrzał wznoszący się atomowy grzyb, no znalazł grzyba, weź ten dolar, złoty, no чаще weź tego dolara, złotego.*

Полное проведение подобной тенденции в полесских говорах украинского языка приводит к тому, что возникает новое противопоставление мужского одушевленного рода (употребляется только одна форма) и мужского неодушевленного (в винительном могут употребляться обе формы).

#### § 48. РАССТОЯНИЕ МЕЖДУ КАТЕГОРИЯМИ КАК ТИПОЛОГИЧЕСКИЙ ПОКАЗАТЕЛЬ

Определив род через длину  $\Gamma$ - $S$ -цепей (см. § 45), Маркус предложил использовать длины  $\Gamma$ - $S$ -цепей и для измерения расстояния между родами.

**0. 6. 26.** Пусть имеется число  $n$  такое, что 1) любое слово из рода  $P_1$  может быть связано с любым словом из рода  $P_2$  при помощи  $\Gamma$ - $S$ -цепи длины меньшей или равной  $n$  и 2) существуют слова  $x$  из  $P_1$  и  $y$  из  $P_2$  такие, что  $x$  и  $y$  нельзя связать  $\Gamma$ - $S$ -цепью длины меньше  $n$ . Тогда число  $(n-3)$  является расстоянием между родом  $P_1$  и  $P_2$ .

Если при этом такой  $\Gamma$ - $S$ -цепи не существует, то мы условимся говорить, что  $n = \infty$ .

**Примечание.** Наша формулировка несколько отличается от формулировки Маркуса, который берет в качестве расстояния не число  $n-3$ , а просто число  $n$ . Но тогда не выполняется аксиома невырожденности расстояния и сами числа не очень показательны. Предложение изменить определение Маркуса исходит от А. В. Гладкого.

Расстояние, введенное Маркусом, является достаточно удобным типологическим критерием, в особенности когда изучаются различные фрагменты языка.

Во французском языке слова мужского и женского рода принадлежат на любом фрагменте разным разделам и поэтому не

<sup>17</sup> См.: А. Heinz. System przypadkowy języka polskiego. Kraków, 1965, стр. 96.

существует  $n$ , удовлетворяющего определению 6. 26. Таким образом, во французском языке расстояние между родами равно бесконечности. В русском языке на фрагменте, рассмотренном в § 34, стр. 126, расстояние между любыми двумя родами равно единице. Эту характеристику он разделяет с соответствующими фрагментами немецкого и латинского языков. В сербском языке на рассматриваемом фрагменте расстояние между мужским и женским родами, так же как и во французском языке, равно бесконечности, если не включать в рассмотрение слова типа *очи* (ср. стр. 167), и трем, если включить соответствующие слова. В последнем случае положение в сербском такое же, как и в румынском. До сих пор не найдено других типов расстояний, чем  $n=0$ ,  $n=1$ ,  $n=3$  и  $n=\infty$ . Это интересно как показатель того, как мало в естественных языках числовых градаций. Однако такой минимальный диапазон возможностей снижает ценность этой меры для типологии.

Другой недостаток предложенной меры расстояния в том, что она непосредственно связана с категорией рода и не может быть применена к другим категориям. Последнее, однако, существенно, ибо мы хотим иметь возможность сравнивать показатели, полученные для разных категорий.

Поэтому мы введем другую меру расстояния между категориями.

0. 6. 27. Выберем две категории  $x_1$  и  $x_2$ , определяемые одной категориальной функцией (т. е. две категории казуса или две категории генуса), и составим всевозможные пары категориальных пучков, различающихся только этими категориями (пусть

Позиция	Пара падежей											
	И/Р <sub>1</sub>	И/Р <sub>2</sub>	И/Д	И/В	И/Тв	И/Пр <sub>1</sub>	И/Пр <sub>2</sub>	Р <sub>1</sub> /Р <sub>2</sub>	Д/Р <sub>2</sub>	Р <sub>1</sub> /Д	Р <sub>1</sub> /В	В/Р <sub>2</sub>
1. МО, ед. ч.	+	+	+	+	+	+	+	-	+	+	+	+
2. МНО, ед. ч.	+	+	+	-	+	+	+	+	-	+	-	-
3. ЖО, ед. ч.	+	+	+	+	+	+	+	+	+	+	+	+
4. ЖНО, ед. ч.	+	+	+	+	+	+	+	-	+	+	+	+
5. СО, ед. ч.	+	+	+	-	+	+	+	-	+	+	+	+
6. СНО, ед. ч.	+	+	+	-	+	+	+	-	+	+	+	+
7. МО, мн. ч.	+	+	+	+	+	+	+	-	+	+	-	-
8. МНО, мн. ч.	+	+	+	-	+	+	+	-	+	+	+	+
9. ЖО, мн. ч.	+	+	+	+	+	+	+	-	+	+	-	-
10. ЖНО, мн. ч.	+	+	+	-	+	+	+	-	+	+	+	+
11. СО, мн. ч.	+	+	+	+	+	+	+	-	+	+	-	-
12. СНО, мн. ч.	+	+	+	-	+	+	+	-	+	+	+	+
Расстояние	1	1	1	6/12	1	1	1	1/12	11/12	1	8/12	8/12



Мы используем этот метод определения расстояния прежде всего для падежей. Начнем с русского языка (таблица на стр. 170—171).

Таким образом, здесь интересно то, что, если для большинства пар расстояние более  $1/2$ , то между двумя родительными расстояние равно  $1/12$ ; между двумя предложными расстояние равно  $2/12$  (а если не учитывать разницу ударений в формах типа *ночи*, то также  $1/12$ ), что, по-видимому, и объясняет обычное неразграничение этих падежей в грамматиках.

Для других языков мы не будем выписывать всех пар, а только те, которые дают сильные отклонения от единицы. В чешском языке интересны следующие пары падежей.

Позиция	Пара падежей			
	И/В	И/Зв	Д/М	В/Зв
1. МО, ед. ч.	+	+	—	+
2. МНО, ед. ч.	—	+	+	+
3. Ж, ед. ч.	+	+	—	+
4. С, ед. ч.	—	—	+	—
5. МО, мн. ч.	+	—	+	+
6. МНО, мн. ч.	—	—	+	—
7. Ж, мн. ч.	—	—	+	—
8. С, мн. ч.	—	—	+	—
Расстояние	$3/8$	$3/8$	$6/8$	$4/8$

Здесь интересно, что противопоставление И и В несколько ослаблено по сравнению с русским языком. Слабо противопоставлены И и Зв (а также В и Зв). Намечается некоторое ослабление противопоставления Д и М.

**Примечание.** Мы отвлекаемся от различия, состоящего в том, что в некоторых типах только дательный, но не местный имеет факультативный вариант с *-ovi*, а также от субстантивированных прилагательных. Нам важно сейчас, что в большинстве слов мужского одушевленного рода местный и дательный не различаются в единственном числе. Если же учитывать указанные явления, то все равно будет иметь место ослабление противопоставления Д и М, хотя и меньшее, а именно в этом случае  $(Д/М) = 7/8$  (полное совпадение лишь в женском роде).

В сербском языке для большинства пар падежей расстояние равно единице. Исключения составляют пары, представленные в таблице (стр. 173).

Особенно интересно сопоставить эти данные с чешским языком, где мы имеем тот же набор падежей и тот же набор позиций.

Позиция	Пара падежей					
	И/В	И/Зв	Д/Тв	Д/М	В/Зв	М/Тв
1. МО, ед. ч.	+	+	+	—	+	+
2. МНО, ед. ч.	—	+	+	+	+	+
3. Ж, ед. ч.	+	+	+	—	+	+
4. С, ед. ч.	—	—	+	—	—	+
5. МО, мн. ч.	+	—	—	—	+	—
6. МНО, мн. ч.	+	—	—	—	+	—
7. Ж, мн. ч.	—	—	—	—	—	—
8. С, мн. ч.	—	—	—	—	—	—
Расстояние	4/8	3/8	4/8	1/8	5/8	4/8

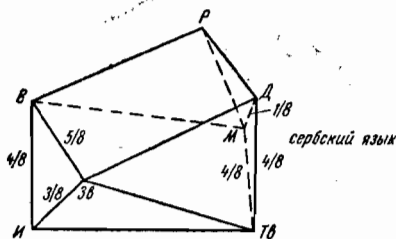
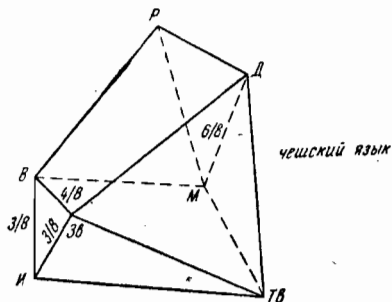
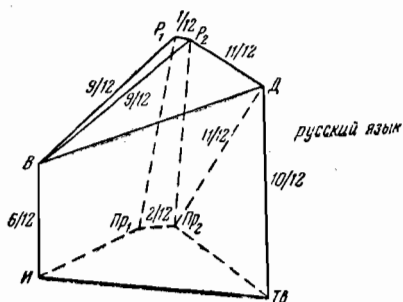
Если противопоставление И и В здесь несколько укреплено по сравнению с чешским, то резко ослабилось противопоставление Д и М, практически это один падеж, так как в единственной позиции, где они различаются, разница сводится лишь к ударению. Следует констатировать, что резко снизилось и противопоставление М и Т. По сильно противопоставленным падежам сербская система приближается к болгарской, что несомненно связано с вхождением обоих языков в балканский союз. К сожалению, для болгарского языка вычисление расстояний затруднено, так как мы определяли там падежи в первую очередь на основе безударных местоимений, а здесь исследуются словоформы существительных. Обращение к языкам типа болгарского несомненно обнаруживает ограниченную применимость (и значимость) модели расстояния и наводит на мысль о желательности такого ее видоизменения, при котором каждому индексу придается определенный вес в зависимости от числа словоформ (или типов словоформ), для которых имеется различие между падежами в данной позиции<sup>18</sup>. К сожалению, мы не располагаем соответствующими статистическими данными по всем языкам и поэтому нет смысла вводить подобную формулу.

Мы видели, что для большинства языков и приведенная более грубая формула приводит к результатам, которые значимы с типологической точки зрения.

Изобразим теперь часть из приведенных данных в виде схем, наглядно показывающих отношения между падежами в некоторых славянских языках.

<sup>18</sup> Идея учета числа типоформ, различающих две падежные функции, была успешно использована Т. В. Цивьян («Имя существительное в балканских языках». М., 1965).

Эти схемы наглядно подтверждают тенденцию к стиранию противопоставления между дательным и местным, о которой говорилось выше и о тенденции сербского языка к четырехпадежной системе (стр. 176).



Для родов мы не будем выписывать таблицы, так как это заняло бы слишком много места, а представим лишь расстояния, вычисленные по той же формуле О. Г. Карпинской<sup>19</sup> (таблица, стр. 175).

Теоретически наибольшее расстояние между двумя родами равно 1 (когда во всех падежах они имеют разные согласования), в славянских языках — 10/14 ( $\rho$  (Ж/С) в сербском и чешском,  $\rho$  (Ж/ЛМ) в польском языке). Наименьшее расстояние (теоретически оно равно нулю) в славянских языках составляет 1/14 ( $\rho$  (МО/МНО) в сербском,  $\rho$  (НМО/НЛМО) в польском) и соответственно 1/12 ( $\rho$  (СО/СНО) в русском языке) так как эти пары различаются одной формой.

Во всех славянских языках расстояния между средними и мужскими родами меньше, чем между этими родами и женскими. Но конкретное их расположение друг относительно друга вполне специфично и зависит, в частности, от качества и числа родов данной системы. Относительно равномерное отстояние родов представлено в сербском языке ( $\rho$  (Ж/С)=10/14,  $\rho$  (Ж/МНО)=

<sup>19</sup> Последняя часть параграфа написана по просьбе автора О. Г. Карпинской; см. также: О. Г. Карпинская. Методы типологического описания. . . , стр. 92—95.

Род	Язык				Род	Язык			
	польский	русский	сербский	чешский		польский	русский	сербский	чешский
JM/НЛМО	3/14				MO/ЖO		6/12		
LM/МНО	3/14				MO/CO		2/12		
МНО/Ж	7/14		9/14	7/14	MO/CHO		3/12		
LM/Ж	10/14				МНО/ЖO		7/12		
LM/С	6/14				МНО/ЖНО		6/12		
НЛМО/МНО	1/14				МНО/CO		3/12		
НЛМО/Ж	7/14				МНО/CHO		2/12		
НЛМО/С	3/14				MO/ЖНО		7/12		
МНО/С	3/14		6/14	6/14	ЖO/CO		6/12		
Ж/С	7/14		10/14	10/14	ЖO/CHO		7/12		
MO/МНО		2/12	1/14	3/14	ЖНО/CO		7/12		
MO/Ж			9/14	9/14	ЖНО/CHO		6/12		
MO/С			6/14	6/14	CO/CHO		1/12		
					MO/ЖНО		7/12		

$=\rho(\text{Ж/МО})=9/14$ ,  $\rho(\text{С/МО})=\rho(\text{С/МНО})=6/14$ , причем женский и средний роды находятся на одинаковом расстоянии и от мужского одушевленного и от мужского неодушевленного родов — в то же время расстояние между этими последними минимально:  $\rho(\text{МО/МНО})=1/14$ .

В чешском языке сохраняется то же расстояние между женским и средним родами, но мужской одушевленный и мужской неодушевленный занимают относительно женского и среднего родов почти симметричное положение  $\rho(\text{Ж/МО})=\rho(\text{Ж/МНО})=7/14$ ,  $\rho(\text{С/МО})=\rho(\text{С/МНО})=6/14$ . Мужской одушевленный чешский резко отодвинут по сравнению с сербским от мужского неодушевленного —  $\rho(\text{МО/МНО})$  в чешском равно  $3/14$ .

Сербский и чешский языки имеют одинаковые наборы родов. В верхнедунайском и польском языках, где, кроме женского, среднего и мужского неодушевленного, представлены еще лично-мужской и нелично-мужской одушевленные роды, значительно сокращено расстояние между женским и средним родами:  $\rho(\text{Ж/С})=7/14$ . В этих языках, также в отличие от чешского и сербского, мужские роды по-разному отстают от среднего:  $\rho(\text{С/ЛМ})=6/14$ ,  $\rho(\text{С/МНО})=\rho(\text{С/НЛМО})=3/14$ .

В русском языке представлены не только мужской одушевленный и мужской неодушевленный, но также одушевленные и неодушевленные женские и средние роды. При этом, если в поль-



ском и верхнелужицком языках по сравнению с сербским и чешским сокращено расстояние только между средним и мужским неодушевленным, в русском языке расстояние между средними и мужскими родами минимум в два раза меньше, чем между ними и женскими.

В расположении родов друг относительно друга можно видеть действие разных тенденций. С одной стороны, здесь возникает вопрос о противопоставлении женской родовой области неженской, который много раз поднимался для славянских языков<sup>20</sup>. Как это видно, данное противопоставление, подкрепляемое и семантическими признаками<sup>21</sup>; в первую очередь относится к русскому языку, нейтрально в отношении чешского и скорее нерелевантно для сербского. Наибольшее различие русского и сербского языков в данном случае объясняется тем, что в русском языке во множественном числе нейтрализуются роды во всех падежах, а в сербском языке — только в части. Если принять во внимание, что между мужским одушевленным и мужским неодушевленным родами в сербском языке представлено наименьшее в славянских языках расстояние (1/14), эту систему можно рассматривать как наиболее приближающуюся к общеславянской трехродовой в первоначальный период возникновения одушевленности.

В польском и верхнелужицком языках противопоставление женского рода неженским во многом оказывается перекрытым новыми тенденциями, связанными с функционированием лично-мужских родов.

С выделением лично-мужского и нелично-мужского одушевленного родов мужской неодушевленный род потерял свой семантический признак, точнее говоря, он по своей семантике максимально сблизился со средним родом. Это сближение произошло и на формальном уровне, так что мужской неодушевленный род в современных польском и верхнелужицком языках занимает симметричное положение между средним и лично-мужскими родами  $\rho(\text{МНО}/\text{С}) = \rho(\text{МНО}/\text{ЛМ}) = 3/14$ . С другой стороны, женский род также максимально среди славянских языков отстоит от лично-мужского рода ( $\rho(\text{Ж}/\text{ЛМ}) = 10/14$ ). Таким образом, есть основания для польского и верхнелужицкого языков говорить о противопоставленности лично-мужской родовой области нелично-мужской<sup>22</sup>.

<sup>20</sup> См.: Р. Нахтигаль. Славянские языки. М., 1963; А. М. Schenker. Gender categories in Polish. «Language», v. 31, № 3.

<sup>21</sup> R. Jakobson. The gender pattern of Russian. «Studii și cercetări lingvistice. Omagiu lui Al. Graur, XI. București, 1960.

<sup>22</sup> Недаром во множественном числе в польском языке иногда выделяют лично-мужской и нелично-мужской «вещественный» род, см.: О. Г. Карпинская. Указ. соч.

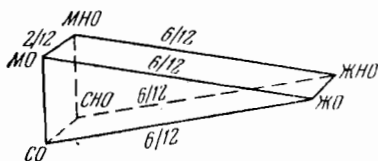
В русском языке наличие шестиродовой системы также наложило отпечаток на общее противопоставление женских родов неженским. Это выражается в том, что расстояние отдельно внутри одушевленных и отдельно внутри неодушевленных меньше, чем между одушевленным и неодушевленным родами, относящимися к разным включающим родам:

$$\begin{aligned} p(\text{МО}/\text{ЖНО}) &= p(\text{МНО}/\text{ЖНО}) = p(\text{ЖО}/\text{СО}) = \\ &= p(\text{ЖНО}/\text{СНО}) = 6/12 \end{aligned}$$

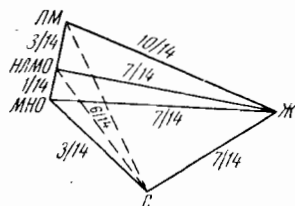
и в то же время

$$\begin{aligned} p(\text{МО}/\text{ЖНО}) &= p(\text{МНО}/\text{ЖО}) = p(\text{ЖО}/\text{СНО}) = \\ &= p(\text{ЖНО}/\text{СО}) = 7/12 \end{aligned}$$

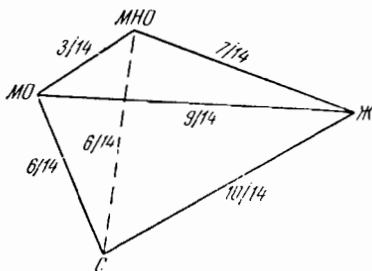
русский



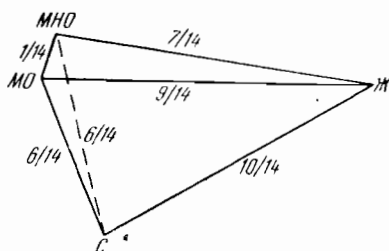
польский, верхнепольский



чешский



сербский



## НЕКОТОРЫЕ ВОПРОСЫ, СВЯЗАННЫЕ С СИНТАГМАТИЧЕСКИМИ МОДЕЛЯМИ

### § 49. ЛИНГВИСТИЧЕСКИЙ АСПЕКТ ГРАММАТИК С КОНЕЧНЫМ ЧИСЛОМ СОСТОЯНИЙ

В § 20 мы определили грамматику с конечным числом состояний как грамматику, каждое правило которой имеет вид  $A \rightarrow aB$ ,  $A \rightarrow \Lambda$  или  $A \rightarrow B$ , где  $a$  входит в основной словарь  $V$ ,  $\Lambda$  есть пустое слово, а  $A$  и  $B$  принадлежат вспомогательному словарю  $V_1$ . В этом вспомогательном словаре выделен специальный символ  $A_0$ , который мы будем называть начальным (его роль аналогична роли символа «предложение» в грамматиках непосредственно составляющих). Остальные символы из  $V_1$  интерпретируются (согласно условию, изложенному в § 20) как имена грамматических категорий<sup>1</sup>, в частности как имена категориальных пучков.

Условно присоединим к основному словарю пустую цепочку  $\Lambda$ . Тогда любое правило может быть представлено в виде  $A \rightarrow aB$  или  $A \rightarrow a$ , где  $a \in V$  или  $a = \Lambda$ .

Вывод в грамматике с конечным числом состояний (см. определение 3.9) имеет очень простой вид, а именно  $(A_0, a_1A_1, a_1a_2A_2, \dots, a_1a_2 \dots a_{n-1}A_{n-1}, a_1a_2 \dots a_{n-1}a_nA_0, a_1a_2 \dots a_{n-1}a_n)$ , где  $A_0 \rightarrow a_1A_1, A_1 \rightarrow a_2A_2, \dots, A_{n-1} \rightarrow a_nA_n, A_n \rightarrow \Lambda$  суть правила грамматики. Символ  $A_0$  соответствует символу «Предложение» в определении 3.8. Заметим, что в любой цепочке, выводимой из  $A_0$ , может быть не более одного символа из  $V_1$ . Фразой, производимой грамматикой с конечным числом состояний, является любая цепочка слов основного словаря  $V$ , заканчивающая вывод.

Объясним теперь происхождение термина «грамматика с конечным числом состояний».

0.7.1. Пусть  $G$  — некоторая грамматика с конечным числом состояний. Поскольку все правила этой грамматики (при включении пустого слова  $\Lambda$  в  $V$ ) имеют вид  $A \rightarrow aB$ , то ей можно поставить в соответствие определенный граф  $G^*$ , построенный следующим образом:

<sup>1</sup> Эта интерпретация восходит к работе: Y. Bar-Hillel, E. Shamir, *Finite state languages: formal representations and adequacy problems*, «Bull. of the research council of Israel», Т. 8, 1960.

а) множество элементов вспомогательного словаря есть множество вершин графа;

б) если  $A \rightarrow aB$  есть правила грамматики, то ему соответствует ребро графа  $\langle A, B \rangle$ , которое помечается символом  $a$ ;

в) правилу  $A \rightarrow \Lambda$  соответствует ребро  $\langle A, A_0 \rangle$ , помеченное  $\Lambda$ ;

г) других ребер у графа нет.

Граф, соответствующий грамматике, мы будем называть диаграммой грамматики.

Пусть, например, грамматика имеет вид:

$$A_0 \rightarrow aB$$

$$A_0 \rightarrow bB$$

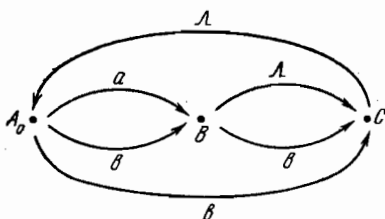
$$A_0 \rightarrow bC$$

$$C \rightarrow \Lambda$$

$$B \rightarrow \Lambda C$$

$$B \rightarrow bC$$

Она производит, в частности, цепочки:  $a, b, ab, bb, abab, bbbab$  и т. п. Данной грамматике соответствует граф следующего вида:



Такие диаграммы можно интерпретировать как устройства, принимающие  $n$  состояний (каждое состояние соответствует вершине графа). Если из состояния  $A$  возможен переход в состояние  $B$ , то пара  $\langle A, B \rangle$  есть ребро графа. Если при переходе из  $A$  в  $B$  может выдаваться слово  $a$ , тогда ребро помечено этим словом. Порождение фразы можно представить как прохождение этим устройством через ряд состояний  $C_0, C_1, \dots, C_n, C_0$ , начиная с начального и кончая им.

Рассмотрим теперь некоторые лингвистически важные виды грамматик с конечным числом состояний, или, как мы будем говорить, следуя А. В. Гладкому<sup>2</sup>, автоматных грамматик.

**0.7.2.** Автоматная грамматика называется однозначной, если в ней для любых состояний  $C_i, C_j$  и  $C_k$  и для любого непустого слова  $x$  при наличии правил  $C_i \rightarrow xC_j$  и  $C_i \rightarrow xC_k$  имеет место  $C_j = C_k$ .

<sup>2</sup> А. В. Г л а д к и й. Конфигурационные характеристики языков. «Проблемы кибернетики», вып. 10. М., 1963, стр. 251—260.

Иначе говоря, в однозначной автоматной грамматике при возможности перехода в два разных состояния не может производиться одно и то же непустое слово.

**0.7.3.** Автоматная грамматика называется грамматикой без омонимии<sup>3</sup>, если в ней для любых состояний  $C_i$  и  $C_j$  и для любого непустого слова  $x$  из наличия правил  $C_i \rightarrow xC_k$  и  $C_j \rightarrow xC_l$  следует  $C_i = C_j$ .

Назовем **ограниченной** грамматикой однозначную автоматную грамматику без омонимии.

Можно было бы показать, что грамматика является ограниченной тогда и только тогда, когда ее диаграмма не содержит двух разных ребер, помеченных одним и тем же непустым словом.

Модель ограниченной грамматики накладывает, разумеется, очень жесткие требования на язык, однако она позволяет поставить некоторые интересные вопросы. Первый из них состоит в следующем. Мы разрешили, чтобы при некоторых переходах выдавались пустые слова. Это, вообще говоря, приводит к тому, что любую фразу можно производить разнообразными способами, вставляя сколь угодно много промежуточных «нулевых переходов», т. е. состояний, после которых производится пустое слово. Поэтому возникает вопрос о выделении тех позиций, в которых целесообразно иметь нулевые переходы. Уже Хомский и Миллер<sup>4</sup> выделили следующий случай: они разбирали класс грамматик, в котором пустое слово выдается каждый раз, когда заканчивается фраза. При этом оказывается, что множество языков, порождаемых такими грамматиками, тождественно множеству языков, порождаемых грамматиками, диаграммы которых не содержат стрелок, помеченных пустым словом.

Нулевое слово (ср. знак  $\neq$  в морфонологии) может рассматриваться как знак конца предложения. Поэтому для любой автоматной грамматики может быть построена эквивалентная ей, но такая, в которой одна фраза отделяется от другой. Именно такие грамматики мы и будем рассматривать в дальнейшем.

Теперь мы хотим показать, что в ограниченных грамматиках часто возникает необходимость введения нулевых переходов в связи с некоторыми внутренними свойствами порождаемых фрагментов.

Рассмотрим язык  $L_1$ , в котором реальны фразы вида  $ca^nb$ ,  $cd^me$ , где  $n, m = 0, 1, \dots$  и запрещены цепочки  $cdb$  и  $cae$ .

**Т. 7.1.** Не существует ограниченной грамматики, совместимой с языком  $L_1$ , и такой, что в ней нет нулевых переходов в начальные состояния.

<sup>3</sup> Это понятие введено А. В. Гладким (см. указ. соч.). Там же анализируются некоторые важные свойства грамматик без омонимии.

<sup>4</sup> Н. Хомский, Г. Миллер. Языки с конечным числом состояний. «Кибернетический сборник», 4, 1962.

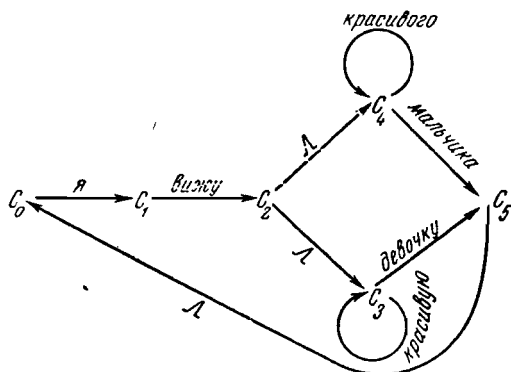
Иначе говоря, существуют языки, которые могут быть порождены ограниченной грамматикой только при условии выдачи пустых слов при переходе в некоторые неначальные состояния, т. е. еще не доходя до конца фразы. Интересно, что язык  $L_1$  изоморфен важным фрагментам естественных языков.

Возьмем множество реальных фраз (7. 2):

*я вижу мальчика; я вижу красивого мальчика; я вижу красивого ... красивого мальчика; я вижу девочку; я вижу красивую девочку; я вижу красивую ... красивую девочку;*

при запрещенных цепочках: *\*я вижу красивого девочку; \*я вижу красивую мальчика.*

Ограниченная грамматика, совместимая с этим фрагментом и отграничивающая одну фразу от другой, имеет вид:



Иначе говоря, система правил выглядит так:

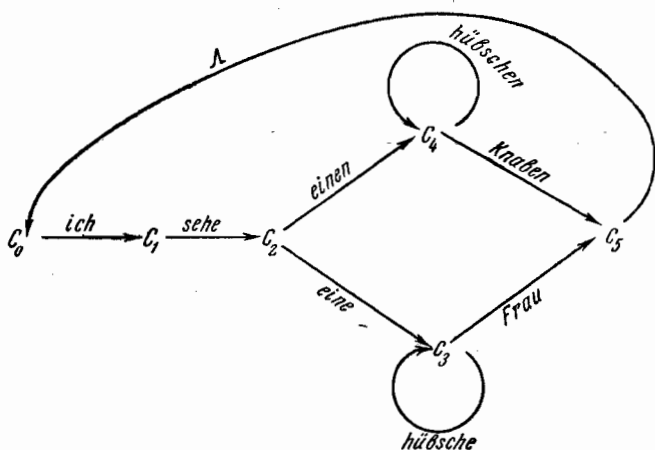
$$\begin{aligned}
 C_0 &\rightarrow \text{я } C_1 \\
 C_1 &\rightarrow \text{вижу } C_2 \\
 C_2 &\rightarrow \Lambda C_4 \\
 C_2 &\rightarrow \Lambda C_3 \\
 C_4 &\rightarrow \text{красивого } C_4 \\
 C_3 &\rightarrow \text{красивую } C_3 \\
 C_4 &\rightarrow \text{мальчика } C_5 \\
 C_3 &\rightarrow \text{девочку } C_5 \\
 C_5 &\rightarrow \Lambda C_0
 \end{aligned}$$

Здесь  $C_1$  может интерпретироваться как категориальный пучок: (1-е л. ед. ч., наст., несов., переход.),  $C_4$  как категориальный пучок (МО ед. ч., В),  $C_3$  как (ЖО, ед. ч., В).

Что касается состояния  $C_2$ , при выходе из которого порождаются пустые слова, то его можно интерпретировать как категорию дополнения, т. е. более абстрактную категорию члена предложения. Это показывает, что грамматика очень простая

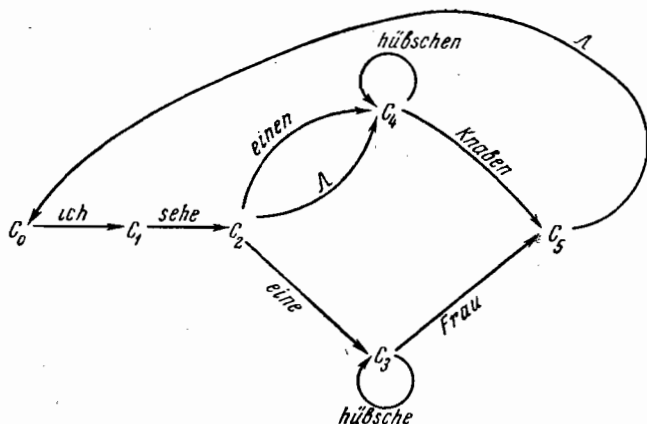
строения может при наложении разумных дополнительных ограничений помочь вскрыть существенные лингвистические свойства.

Можно заметить одно важное типологическое различие между языками, имеющими препозитивный артикль, и языками без артиклей. Так, в немецком языке для порождения фраз *ich sah eine hübsche Frau, ich sehe einen hübschen Knaben* можно обойтись диаграммой без пустых слов:



И только если одновременно нужно произвести фразу с нулевым артиклем: *ich sehe hübschen . . . hübschen Knaben* обязательно n раз

появляется пустое слово:

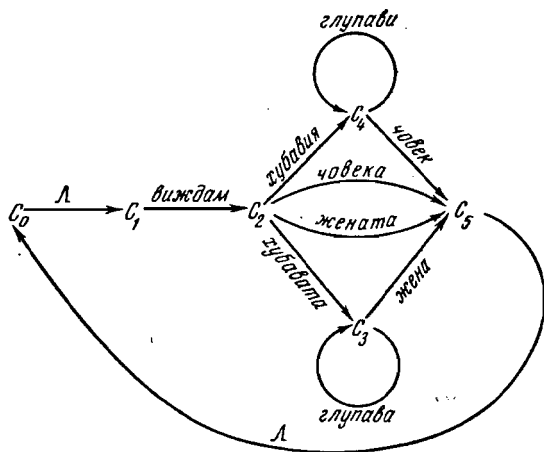


Здесь любопытно не только то, что «нулевому артиклю» обязательно соответствует пустое слово на диаграмме, но и то, что

в языках типа русского пустое слово играет роль артикля на диаграмме. Болгарский язык дает здесь (в фрагментах, где присутствует определенный артикль) диаграммы, отличные как от русского языка, не имеющего артикля, так и от немецкого, имеющего препозитивный артикль.

Возьмем почти тот же фрагмент, что и соответствующий русский. Реальные фразы: *виждам жената*; *виждам хубавата жена*; *виждам хубавата глупава жена*; *виждам човека*; *виждам хубавия човек*; *виждам хубавия глупави човек*;

при запрещенных цепочках: \**виждам хубавата жената*; \**виждам хубавата, глупавата жена*; \**виждам хубавия човека*; \**виждам хубавия глупавия човек*. Этот фрагмент порождается схемой:



Эта схема интересна тем, что в ней место пустого слова в русских схемах и артикля в немецких занимает членная форма прилагательного. Причина этого в том, что в болгарском языке — в отличие от других славянских — сочетания «прилагательное + существительное» не являются регулярными сегментами (ср. определение 3. 6).

Легко заметить, что определители регулярного сегмента представляются на диаграммах циклами, т. е. правилами вида  $C_1 \rightarrow xC_1$ .

Если исходить из нашего принципа, по которому каждое состояние соответствует категориальному пучку, то тогда мы получаем важное следствие, согласно которому определители регулярных сегментов не имеют на диаграмме собственного категориального пучка, а место их целиком определяется категориальным пучком ядра сегмента. Таким образом, в случае определителей регулярных сегментов диаграмма довольно хорошо отражает синтаксическую зависимость, в то время как в других слу-



чаях из диаграммы невозможно заключить, какое слово является главным, а какое зависимым.

Указанные свойства можно использовать и для экономного восстановления порождающего процесса по заданному множеству реальных и запрещенных цепочек, если установлено, что некоторое сочетание  $xu$  является регулярным сегментом с ядром  $y$ , то правила строятся следующим образом:  $C_1 \rightarrow xC_1$ ,  $C_1 \rightarrow yC_2$ , а в остальных случаях правила строятся на основании линейного порядка: если встретилось сочетание  $xu$ , то обязательно вводится правило  $C_1 \rightarrow xC_2$ ,  $C_2 \rightarrow yC_3$ . Интересной проблемой является нахождение общего способа построения минимальной схемы для ограниченного языка (включая и вопрос о введении нулевых переходов).

Вернемся к лингвистическому анализу языка, порожденного ограниченной грамматикой.

Понятие ограниченной грамматики представляет интерес и в связи с понятием рамки:

**0.7.4.** Реальная фраза  $f = x_1 \dots x_n$  называется фразой с рамкой  $(i, j)$ , если в нее входят слова  $x_i$  и  $x_j$  ( $j \geq i + 2$ ), для которых найдутся слова  $y_i$  и  $y_j$ , такие, что цепочка  $g = x_1 \dots x_{i-1} y_i x_{i+1} \dots x_j x_{j+1} \dots x_n$  запрещена, в то время как существует реальная фраза  $g' = x_1 \dots x_{i-1} y_i x_{i+1} \dots x_{j-1} y_j x_{j+1} \dots x_n$ .

Иначе говоря, фраза с рамкой — это такая фраза, в которой замена  $x_i$  на  $y_i$  требует замены  $x_j$  на  $y_j$ , причем  $x_j$  расположено дистантно от  $x_i$ .

**Т. 7. 2.** Язык, в котором имеется хотя бы одна фраза с рамкой  $(i, j)$ , не совместим с ограниченной грамматикой.

В славянских языках рамки возникают прежде всего из-за свободного порядка слов, например *Она без мысли и без речи | На том смеется берегу* (А. Блок). Здесь рамки образуют такие пары слов, как *она* и *смеется*, *там* и *берегу*. Однако в отличие, например, от немецкого языка в славянских языках для каждого предложения с рамкой найдется такое, в котором рамка отсутствует. Это обстоятельство делает полезным следующее определение.

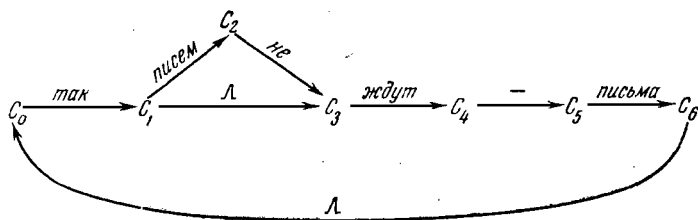
**0.7.5.** Рамка  $(i, j)$  называется у с т р а н и м о й в фразе  $f$ , если найдется реальная фраза  $f_1$ , а) не имеющая рамки, б) составленная из тех же слов и в) тождественная фразе  $f$  по смыслу.

Неустранимые рамки возникают в славянских языках, по-видимому, лишь в сложных предложениях, ср. *если — то, поскольку — постольку* и т. п.

Однако подъязык, включающий такие рамки, относится скорее к книжной научной речи. Вполне возможно, что модели для книжной речи должны выглядеть по существу иначе, чем модели для разговорной речи. С другой стороны, можно было бы допустить, что а) модель порождения состоит из нескольких уров-

ней (ср. § 56 и § 64), б) на уровне простых предложений используется как раз схема ограниченной грамматики.

В пункте б) высказано слишком сильное предположение, правдоподобное лишь в случае частей фраз или простых фраз с обычным порядком слов и с четким разграничением групп подлежащего и сказуемого. В других случаях, хотя формально и можно построить для некоторого множества фраз ограниченную грамматику, она уже не будет иметь никакой объяснительной силы, т. е. потеряет всякую лингвистическую значимость. Например, для фраз: *так писем не ждут* | *так ждут* — *письма* (М. Цветаева) диаграмма, в которой слово *ждут* не производится в разных местах, должна иметь вид:

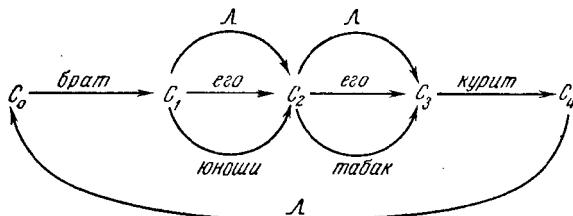


Эта диаграмма не только не объясняет синтаксической структуры обеих фраз, но порождает к тому же вряд ли осмысленную цепочку: *так писем не ждут — письма*.

Разумеется, там, где фразы, которые мы хотим породить, содержат омонимичные слова, вообще вполне естественно отказаться от требования неомонимичности грамматики.

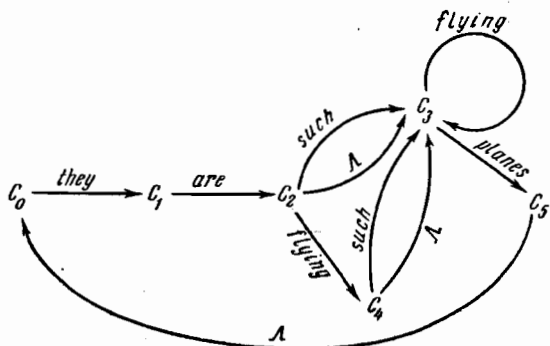
Покажем теперь, что и при омонимичности грамматики понятие пустого слова с лингвистической точки зрения может оказаться весьма полезным.

Рассмотрим предложение *брат его курит*. Омонимичность этого предложения лучше всего вскрывается, если параллельно рассмотреть предложения *брат юноши курит*; *брат табак курит*; *брат юноши табак курит*. Наиболее естественная диаграмма, совместимая с этим фрагментом, имеет вид:



Таким образом, предложение *брат его курит* выдается двумя способами: *брат его* Λ *курит* и *брат* Λ *его курит*, что и свидетельствует об его омонимичности.

Еще более показателен английский пример, часто приводимый именно для того, чтобы показать, что грамматика с конечным числом состояний не дает возможности разграничить случаи омонимии. Речь идет о предложении *they are flying planes* с двумя возможными переводами: 1) 'они водят самолеты'; 2) 'это — летящие самолеты'. Омонимичность такого предложения лучше всего проявляется на фоне таких неомонимичных фраз: *they are such flying planes* 'это такие летящие самолеты' — *they are flying such planes* 'они водят такие самолеты'. Достаточно естественная однозначная диаграмма с пустым словом в конце имеет вид:



Здесь разное распределение нулевых слов:

*They are  $\Lambda$  flying planes / They are flying  $\Lambda$  planes*

показывает две разные возможности членения предложения и таким образом объясняет его омонимичность.

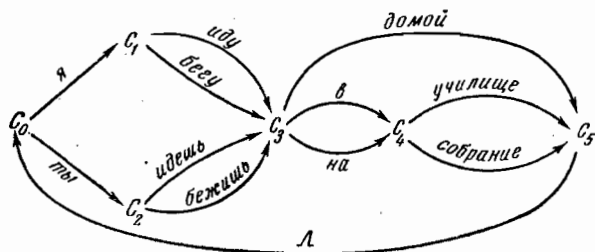
Мы познакомились, таким образом, с двумя лингвистически важными интерпретациями пустого слова. Это, во-первых, знак конца предложения или синтаксической группы и, во-вторых, «нулевой элемент» предложения. Вторая интерпретация особенно существенна. В морфологии понятие нулевой морфемы вполне приобрело права гражданства. В синтаксисе можно ввести аналогично нулевую позицию, которая и вскрывается нашей моделью. Так, в примере *брат его  $\Lambda$  курит* пустое слово может трактоваться как нулевое дополнение при переходном глаголе *курить* (т. е. можно считать, что при таком глаголе дополнение обязательно, но выражено оно может быть нулевым словом).

Таким образом в интерпретации можно было бы разграничить пустые слова как сигналы границы и нулевые слова как члены класса, один из элементов которого обязателен в предложении.

Этот аспект наглядно вскрывает парадигматический принцип, лежащий в основе порождения автоматной грамматикой: одновременное синтезирование сходным способом большого числа предложений сходной структуры.

В этой связи приобретает интерес проведенное нами выше отождествление состояний на схеме с грамматическими категориями, приводящее к тому, что любое состояние становится грамматической характеристикой следующего за ним слова. Но для более последовательного осуществления этого принципа необходимо, по-видимому, ввести некоторые изменения в схему порождения. Рассмотрим следующий пример.

Пусть даны реальные фразы: *я иду домой; я иду в училище; я иду на собрание; ты идешь домой; ты идешь в училище; ты идешь на собрание; я бегу домой; я бегу на собрание, ты бежишь домой; ты бежишь в училище; ты бежишь на собрание.* Естественная схема, насчитывающая шесть состояний, имеет для этих фраз вид:



Здесь можно интерпретировать:

- $C_1$  как категориальный пучок: ед. ч., 1-е л., наст.
- $C_2$  " " " ед. ч., 2-е л., наст.
- $C_4$  " " " В, ед. ч.

При этом, однако, не получают интерпретации в терминах категорий те состояния, из которых стрелки ведут в разные состояния, причем порождаются слова, принадлежащие разным категориальным пучкам, например, состояние  $C_3$ . Поэтому предлагается следующее видоизменение системы правил соответствующей им диаграммы. Пусть в  $S$  входят два правила:

$$C_i \rightarrow xC_k$$

$$C_i \rightarrow yC_j,$$

где  $C_k \neq C_j \neq C_i$ , а  $x$  и  $y$  непустые слова. Тогда мы заменяем эти два правила следующей эквивалентной системой правил

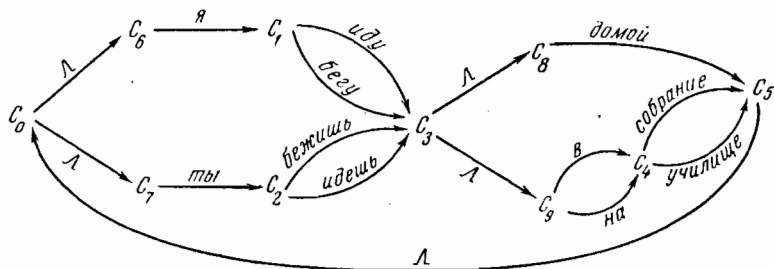
$$C_i \rightarrow \Lambda C_m$$

$$C_i \rightarrow \Lambda C_n$$

$$C_m \rightarrow xC_k$$

$$C_n \rightarrow yC_j$$

В частности, схема принимает тогда вид:

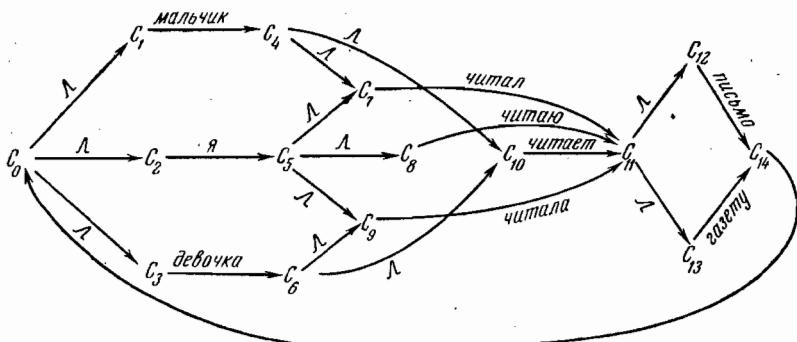


Теперь можно интерпретировать:

- $C_6$  как категориальный пучок: мест., 1-е л., ед. ч.
- $C_7$  » » » мест., 2-е л., ед. ч.
- $C_8$  » » » наречие места
- $C_9$  » » » предлог, управляющий  $V$

Остается найти интерпретацию для тех состояний, при выходе из которых порождается пустое слово. Для того, чтобы нагляднее пояснить лингвистический смысл таких состояний, мы рассмотрим еще порождение фраз: *я читал газету, я читаю газету, мальчик читает газету, девочка читает газету, мальчик читал газету, девочка читала газету, мальчик читает письмо, девочка читает письмо, мальчик читал письмо, девочка читала письмо, я читала газету.*

Диаграмма грамматики, удовлетворяющей, описанному выше условию, может иметь следующий вид:

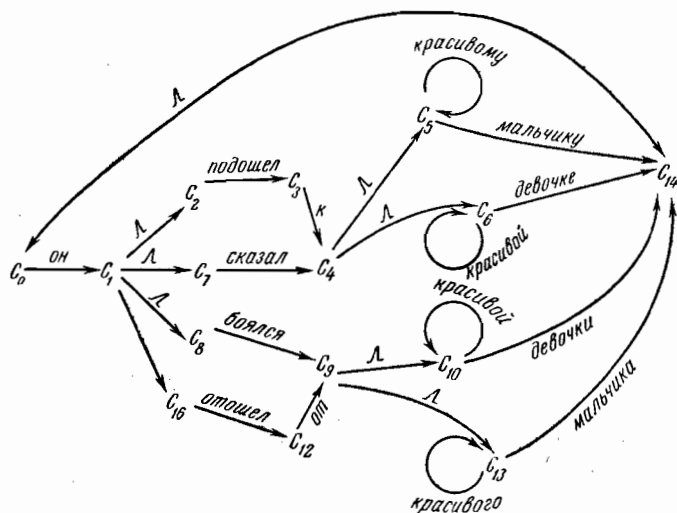


Добавочные состояния (т. е. те, при выходе из которых появляется пустое слово) соответствуют границам между членами предложения традиционной грамматики. Поэтому соответствующие состояния целесообразно интерпретировать как синтаксиче-

ские категории: можно интерпретировать состояния  $C_4$ ,  $C_5$  и  $C_6$  как категории сказуемого и  $C_{11}$  как категорию дополнения. Аналогично на предыдущей схеме  $C_3$  представляет категорию обстоятельства. Нам важны здесь, разумеется, не названия, а сам факт появления подобных синтаксических категорий (ср. на стр. 181 интерпретацию состояния  $C_2$ ).

Итак, при порождении каждой фразы одновременно порождается какое-то количество пустых слов, которые можно интерпретировать как показатели границ синтаксических единств. Посмотрим теперь, какие выводы можно сделать из необходимости или, наоборот, факультативности введения промежуточного состояния и соответственно пустого слова в спорных случаях синтаксического членения фразы.

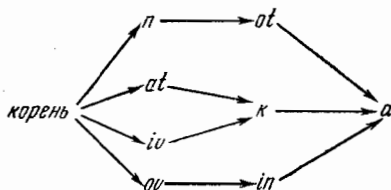
Рассмотрим множество фраз: *он подошел к красивой . . . красивой девочке, он отошел от красивой . . . красивой девочки, он подошел к красивому . . . красивому мальчику, он отошел от красивого . . . красивого мальчика, он боялся красивой . . . красивой девочки, он боялся красивого . . . красивого мальчика, он сказал красивой . . . красивой девочке, он сказал красивому . . . красивому мальчику.* Диаграмма грамматики, порождающей этот фрагмент, имеет вид:



Нас интересует в этой диаграмме порождение групп слов *подошел к* и *отшел от*. Если после слов *к* и *от* введение пустых слов обязательно (в связи с появляющимися далее разными циклами), то перед этими словами промежуточных состояний вводить не требуется, так как, во-первых, между ними не появляются разные циклы и, во-вторых, переход от предыдущего слова к этим словам однозначен.

Таким образом, в этой схеме группы *подошел к* и *отошел от* предстают как единые синтаксические единицы.

Мы увидим дальше, что на уровне синтаксиса нет большой разницы между отдельными словами и синтаксическими группами. С другой стороны, если брать в качестве исходных элементов не словоформы, а морфемы (а в качестве фраз — словоформы), то все явления словообразования и формообразования могут быть объяснены на базе автоматной грамматики. Так, словацкий лингвист Хорецкий изложил все типы морфемной структуры слова в словацком языке в виде схем типа:



Эта схема построена в результате анализа слов: *mastnota*, *šatovka*, *hlasivka*, *rohátka*, *slamienka*, *drevovina*, *plesnivina*, *ladnička* <sup>5</sup>.

Все подобные схемы могут быть интерпретированы как диаграммы довольно простой автоматной грамматики (однозначной и не содержащей циклов, вообще говоря, неомонимичной).

Теоретически вопрос о применении автоматной грамматики к морфологическому анализу был исследован группой французских ученых под руководством математика Б. Вокуа в гренобльском центре по изучению проблем машинного перевода <sup>6</sup>. В этих исследованиях показано, что все процедуры морфологического анализа, применяемые в машинном переводе (в частности, процедура, предложенная И. А. Мельчуком <sup>7</sup>), могут быть истолкованы как специальный вид автоматной грамматики.

Общий подход этой группы к моделированию языка состоит в том, что предлагаемая ими модель содержит два уровня:

а) уровень автоматной грамматики, моделирующий морфологический анализ;

<sup>5</sup> J. Ногеску. *Morfematická štruktúra slovenčiny*. Bratislava, 1964, стр. 86—87. — Легко заметить также, что нет принципиальной разницы между такими схемами и соответствующими схемами на фонологическом уровне, как они разбирались, например, Маркусом и Василиу (см. изложение этих схем: И. И. Ревзин. *Модели языка*, стр. 47—49).

<sup>6</sup> Устное сообщение г. Вокуа в Институте славяноведения; см. также: B. Vauquois, G. Veillon, J. Veurgnes. *Application des grammaires formelles aux modèles linguistiques en traduction automatique*. «Kibernetika», číslo 3, roč. 1. Praha, 1965, стр. 281—289.

<sup>7</sup> И. А. Мельчук. *Морфологический анализ при машинном переводе* (преимущественно на материале русского языка). «Проблемы кибернетики», вып. 6. М., 1961, стр. 207—276.

б) уровень бесконтекстной грамматики, моделирующий синтаксический анализ.

Мы предлагаем расширить сферу применения автоматной грамматики, включив в нее не только анализ и синтез словоформ, но также и анализ и синтез простых словосочетаний (единственная формальная разница между словоформами и словосочетаниями в наших терминах состоит в том, что автоматные грамматики, порождающие словосочетания, будут содержать циклы).

Интересна и аналогия между нашим принципом введения дополнительных состояний для более адекватного лингвистического их истолкования и методом разграничения омонимов, предлагаемым гренобльской группой. В цитированной статье показывается, что если мы хотим получить минимальную схему, синтезирующую достаточное большое число словоформ, то, например, при образовании прилагательных *actif*, *active* и *utile* схема должна содержать правила вида:

$$\begin{aligned}C_0 &\rightarrow \text{acti } C_i \\C_i &\rightarrow vC_k \\C_i &\rightarrow fC_0 \\C_0 &\rightarrow \text{util } C_k \\C_k &\rightarrow eC_0\end{aligned}$$

Построения такой грамматики (фрагмент которой мы привели) достаточно для того, чтобы множество всех словоформ было разрешимым, т. е. для любой цепочки морфем можно было бы сказать, является ли она словоформой данного языка, или нет.

Важно, однако, использовать грамматики для того, чтобы каждой словоформе (или в общем случае — фразе) приписать определенную грамматическую информацию. Это и достигалось в наших грамматиках путем отождествления состояний с грамматическими категориями.

При минимальной диаграмме, фрагмент которой мы привели, *active* и *utile* получают одинаковую грамматическую характеристику, между тем *active* — только женского рода, в то время как *utile* относится как к мужскому, так и к женскому роду. Поэтому в гренобльском центре предлагается, чтобы перед *active* производился специальный символ Ж, а перед *utile* символ М или Ж. Но очевидно, что это эквивалентно введению лишних состояний, ибо Ж и М не являются элементами V (где V интерпретируется как список морфем).

Выше мы упомянули о проблеме разрешимости; поскольку множество словоформ производится автоматной грамматикой без циклов, то оно разрешимо тривиальным способом, ибо оно конечно и любая диаграмма эквивалентна списку, хотя и более удобна, чем список.



Но разрешимым является и любое множество фраз, порождаемое автоматной грамматикой.

Это явствует из следующих соображений. Как показали Хомский и Миллер в цитированной статье, дополнение до языка с конечным числом состояний есть также язык с конечным числом состояний, т. е. множество фраз, составленных из данного словаря и не порождаемых данной автоматной грамматикой, порождается некоторой другой автоматной грамматикой. Таким образом, перечислимым является не только язык с конечным числом состояний, но и его дополнение. Между тем известно, что множество разрешимо, если оно и его дополнение перечислимы.

В заключение этого параграфа заметим следующее. С легкой руки Хомского, очень много сделавшего именно для исследования автоматных грамматик, почти все авторы, пишущие о лингвистических моделях, подчеркивают недостатки этих грамматик, их неадекватность естественным языкам. Мы пытались показать, что некоторые лингвистические аспекты хорошо моделируются автоматными грамматиками, которые с формальной точки зрения имеют ряд несомненных преимуществ (ср. § 64).

## § 50. БЕСКОНТЕКСТНЫЕ ГРАММАТИКИ

Напомним, что по определению 3.12 всякое правило бесконтекстной грамматики непосредственно составляющих имеет вид  $A \rightarrow w$ , где  $A$  есть элемент вспомогательного словаря  $V_1$ , в то время как  $w$  есть непустая цепочка символов над  $V \cup V_1$ .

Символы вспомогательного словаря  $V_1$ , участвующие в правилах, можно классифицировать следующим образом.

**0. 7. 6.** Символ  $A$  называется *н е ц и к л и ч е с к и м*, если не существует таких цепочек  $\chi_1$  и  $\chi_2$  (из которых хотя бы одна непустая), что из  $A$  выводится  $\chi_1 A \chi_2$ . Символ  $A$  называется *ц и к л и ч е с к и м*, если существует хотя бы одна непустая цепочка  $\chi_1$  или  $\chi_2$ , такая, что из  $A$  выводится  $\chi_1 A \chi_2$ .

**П р и м е ч а н и е.** Иногда вместо «циклический» употребляют термин «рекурсивный». Однако этот термин может понадобиться в теории моделей языка в его точном математическом смысле.

**0. 7. 7.** Символ  $A$  называется *л е в о - ц и к л и ч е с к и м*, если существует непустая цепочка  $\chi_1$ , такая, что из  $A$  выводится  $A \chi_2$ , и *п р а в о - ц и к л и ч е с к и м*, если существует непустая цепочка  $\chi_1$ , такая, что из  $A$  выводится  $\chi_1 A$ .

**0. 7. 8.** Символ  $A$  называется *с а м о в с т а в л я ю щ и м с я*, если существуют такие непустые цепочки  $\chi_1$  и  $\chi_2$ , что из  $A$  выводится  $\chi_1 A \chi_2$ .

Из определения 7. 8 следует, что всякий самовставляющийся символ является циклическим, но обратное неверно.

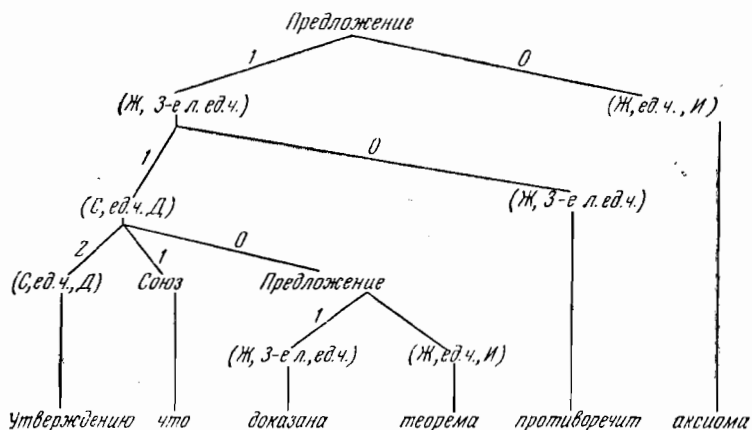
Как доказал Хомский, всякое множество фраз, порождаемое грамматикой без самовставляющихся символов, может быть порождено и автоматной грамматикой.

Фактически самовставляющиеся символы используются лишь при порождении сложных предложений, например таких, как *утверждению, что доказана (еще одна) теорема, противоречила аксиома*. Здесь действительно появляется самовставляющийся символ, как легко убедиться, исходя из следующей системы правил:

1. Предложение  $\rightarrow$  (Ж, 3-е л., ед. ч.) + (Ж, ед. ч., И)
2. (Ж, ед. ч. 3-е л.)  $\rightarrow$  (С, ед. ч., Д) + (Ж, 3-е л., ед. ч.)
3. (С, ед. ч., Д.)  $\rightarrow$  (С, ед. ч., Д) + *что* + предложение
4. (Ж, 3-е л., ед. ч.)  $\rightarrow$  *доказана, противоречила*
5. (Ж, ед. ч., И)  $\rightarrow$  *теорема, аксиома*
6. (С, ед. ч., Д)  $\rightarrow$  *утверждению*

в которой из символа «предложение» выводится последовательность символов: (С, ед. ч. Д) + *что* + предложение + (Ж., 3-е л., ед. ч.) + (Ж, ед. ч., И).

Поскольку теоретически нельзя ограничить число таких самовставлений, а с другой стороны, целесообразно пользоваться именно символом «предложение», то мы имеем дело с подлинным самовставляющимся символом. Это хорошо иллюстрируется следующей схемой:



Легко видеть, что наша грамматика может произвести и такие «сверхсложные» предложения, как *утверждению, что утверждению, что доказана теорема, противоречила аксиома, противоречила аксиома*. Заметим, что такое предложение может оказаться осмысленным.

Вероятно, подобные предложения целесообразно описывать как порождаемые не грамматиками непосредственно составляющих, а грамматиками трансформационными (ср. § 51).

Поэтому если бы задача порождающей грамматики состояла только в том, чтобы а) она определяла все множество правильно построенных фраз (в смысле определения 3.16) и б) разделяла каждое высказывание на сегменты, отмечая в случае омонимичных фраз возможности двоякого членения, то для всех разбираемых нами примеров можно было бы, по-видимому, ограничиться грамматиками с конечным числом состояний.

Однако, как неоднократно подчеркивал почти во всех своих работах Хомский — и в этом огромное достоинство его концепции, — порождающая грамматика должна не только определять множество правильно построенных фраз, но и объяснить понимание структуры каждой фразы говорящим и слушающим. В частности, для понимания предложения необходимо четкое выделение субъектно-предикатной структуры, а это и обеспечивается основной составляющей, маркированной символом «предложение» и распадающейся на две части — субъект и предикат. Интересно, что основная масса остальных составляющих также имеет семантику приписывания признака объекту (или признаку же), но передает эту идею в скрытом, «депредицированном» виде (ср. *маленькая девочка* и *девочка — маленькая*). Именно в передаче субъектно-предикатной структуры семантика непосредственно составляющих, и в этом еще одно соображение в пользу описания сложных предложений, однородных членов и т. п. при помощи других грамматик, а именно грамматик трансформационных.

Указанные содержательные преимущества грамматик непосредственно составляющих для описания целого предложения подкрепляются и их формальными преимуществами. Подобная грамматика естественно приписывает каждому предложению схему вывода, отражающую в известной степени (а именно там, где соответствующие отношения могут быть содержательно истолкованы) систему синтаксических отношений в этом предложении (уточнение этого см. в § 53), причем каждой такой схеме однозначно соответствует определенная скобочная запись, наглядно отражающая членение на составляющие.

Эта особенность облегчает как формальное изучение синтаксических отношений (несколько идеализируя, правда, действительную ситуацию), так и их интерпретацию (основное отличие этих схем от графов состояний в схеме с конечным числом состояний именно в том, что последние гораздо труднее поддаются лингвистической интерпретации).

Разберем теперь на примерах процесс порождения — методом безконтекстной грамматики — фраз, которые в отличие от приведенной выше не содержат самовставляющихся символов, поэтому для облегчения сравнения с предыдущим мы будем разбирать те же фразы, что и в § 49.

Для множества реальных фраз (см. стр. 181) грамматика имеет вид:

1. Предложение  $\rightarrow$  (1-е л., ед. ч., И) + (1-е л., ед. ч., наст. вр. неперех.);

2. (1-е л., ед. ч., наст. вр., неперех.)  $\rightarrow$  (1-е л., ед. ч., наст. вр. перех.) + (М, ед. ч., В) или (Ж, ед. ч., В);

3. (М, ед. ч., В)  $\rightarrow$  (М, ед. ч., В, положит.) + (М, ед. ч., В);

4. (Ж, ед. ч., В)  $\rightarrow$  (Ж, ед. ч., В, положит.) + (Ж, ед. ч., В);

5. (1-е л., ед. ч., И)  $\rightarrow$  я;

6. (1-е л., ед. ч., наст. вр., перех.)  $\rightarrow$  вижу;

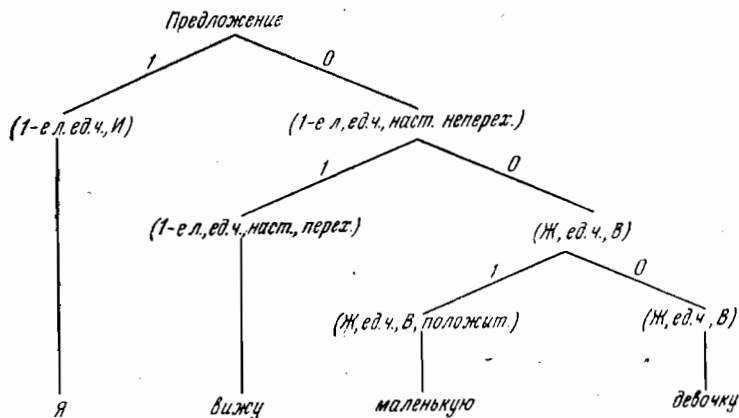
7. (М, ед. ч., В)  $\rightarrow$  мальчика;

8. (Ж, ед. ч., В)  $\rightarrow$  девочку;

9. (М, ед. ч., В, положит.)  $\rightarrow$  маленького;

10. (Ж, ед. ч., В, положит.)  $\rightarrow$  маленькую.

Эта совокупность правил приписывает, например, предложению: Я вижу маленькую девочку следующую схему вывода:]



Правоцикличность символа (Ж, ед. ч., В) наглядно демонстрируется положением и повторением на схеме.

Данному анализу по составляющим однозначно соответствует скобочная запись (я (вижу (маленькую девочку))).

Это верно для всех фраз, порождаемых грамматикой описанного здесь вида.

Интересно посмотреть, как данная грамматика выясняет вопросы омонимии. Рассмотрим правила для порождения того же множества фраз, что и на второй схеме (стр. 185).

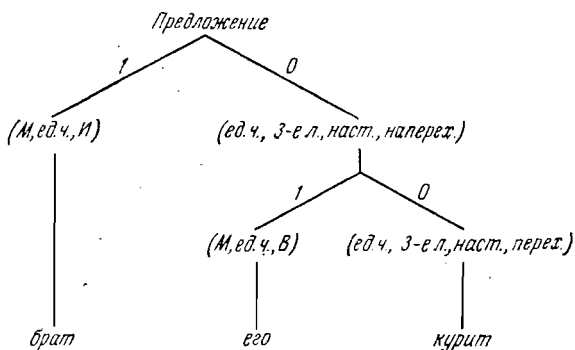
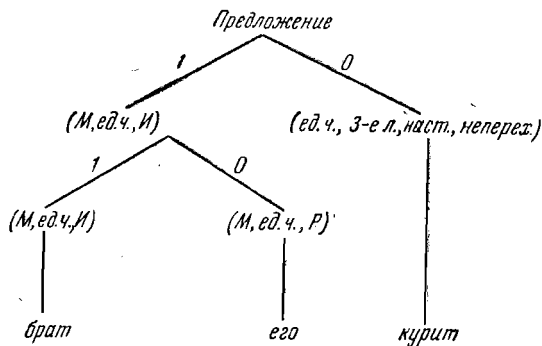
1. Предложение  $\rightarrow$  (М, ед. ч., И) + (ед. ч., 3-е л., наст. вр. неперех.).

2. (М, ед. ч., И)  $\rightarrow$  (М, ед. ч., И) + (М, ед. ч., Р).

3. (М, ед. ч., наст. вр., неперех.)  $\rightarrow$  (М, ед. ч., В) + (М, ед. ч., наст. вр., перех.).

4. (М, ед. ч., И) → *брат*
5. (ед. ч., 3-е л. наст. вр., неперех.) → *курит*
6. (ед. ч., 3-е л. наст. вр., перех.) → *курит*
7. (М, ед. ч., Р) → *его, юноши*
8. (М, ед. ч., В) → *его, табак*

Предложение *брат его курит* имеет две следующие схемы вывода:



В скобочной записи эта неоднозначность отражается следующим образом: а) ((*брат его*) *курит*) и б) (*брат* (*его курит*)).

Иногда об омонимичности (правда, гораздо более слабой) может свидетельствовать и тот факт, что данная фраза может быть произведена той же самой последовательностью правил, но в разном порядке. Возьмем, например, предложение: *Старый товарищ брата приехал*. Его производит следующая совокупность правил:

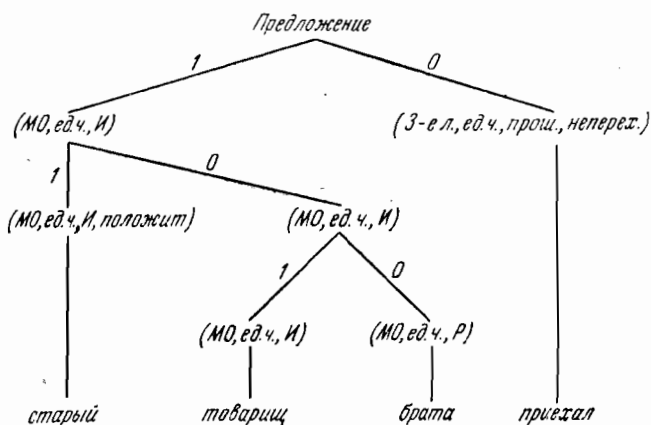
1. Предложение → (МО, ед. ч., И) + (3-е л., ед. ч., прош. вр., М, неперех.).
2. (МО, ед. ч., И) → (М, ед. ч., И, положит.) + (МО, ед. ч., И)
3. (МО, ед. ч., И) → (МО, ед. ч., И) + (М, ед. ч., Р)
4. (МО, ед. ч., И) → *товарищ*

5. (3-е л., ед. ч., прош. вр., неперех.) → *приехал*

6. (МО, ед. ч., И, положит.) → *старый*

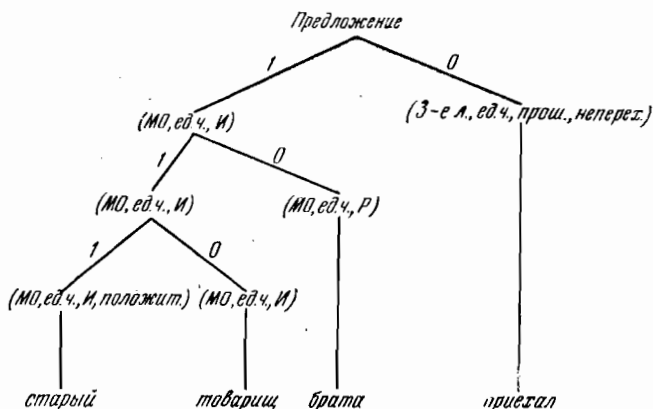
7. (М, ед. ч., Р) → *брата*

Если правила применяются в написанном выше порядке, то мы получаем схему вывода:



Скобочная запись, соответствующая этой схеме, имеет вид:  
((старый (товарищ брата)) приехал).

Если же сначала применять правило 3, а затем правило 2, то схема вывода будет иметь вид:



Скобочная запись, соответствующая этой схеме, имеет вид:  
(((старый товарищ) брата) приехал).

Такая двойственная трактовка однозначных на первый взгляд предложений рассматривается некоторыми авторами как показа-

тель слабости данной модели<sup>8</sup>. С этим вряд ли можно согласиться. Ведь в принципе возможно, хотя практически и не реализуется противопоставление *старый (а не новый) товарищ брата*, с одной стороны, и *старый товарищ брата (а не сестры)*, с другой стороны.

Аналогичен часто приводившийся пример *старые мужчины и женщины*, который может иметь в зависимости от смысла два разных скобочных представления: ((*старые мужчины*) и *женщины*) и (*старые (мужчины и женщины)*)).

Таким образом, подобные примеры не только не свидетельствуют о слабости модели, но наоборот, подчеркивают ее объяснительную силу.

Указанная модель обладает еще одним важным преимуществом: в ней можно оценить сложность структуры выводимого предложения. С одним таким показателем мы уже познакомились — это наличие самовставляющихся символов. Другой показатель связан с понятием глубины предложения, выдвинутым В. Ингве<sup>9</sup>.

0. 7. 9. Рассмотрим произвольную схему вывода предложения. Если из некоторого узла на этой схеме, помеченного нетерминальным символом, выходит  $n$  узлов, то мы припишем каждому из ребер справа налево индексы, идущие от нуля подряд до самого левого ребра, которое ведет в узел, помеченный нетерминальным символом; остальным ребрам приписываются нули. Эта нумерация проводится для всех ребер. Г л у б и н о й символа называется сумма индексов вдоль пути, ведущего от символа «предложения» до этого символа. Глубиной схемы называется максимальная глубина его терминальных символов.

На схемах (см. стр. 193—197) мы поставили индексы на ребрах.

Из этих примеров следует, в частности, что наличие самовставляющихся символов может быть связано с повышением глубины, но что такую же глубину могут иметь и схемы без самовставляющихся символов. Характерно, что разные схемы, соответствующие многозначному предложению, могут иметь и разную глубину.

---

<sup>8</sup> См., например: С. К. Шаумян и П. А. Соболева. Аппликативная порождающая модель и исчисление трансформаций в русском языке. М., 1963, стр. 114. — Также и Курилович считает, что если и существует какое-либо семантическое различие между подобными анализами, то оно «относится к области экспрессии и стиля, но не грамматики» (Е. Курлович. Очерки по лингвистике. М., 1962, стр. 51).

<sup>9</sup> Гипотеза Ингве достаточно хорошо известна, см., например: В. Ингве. Гипотеза глубины. Перев. с англ. «Новое в лингвистике», вып. IV. М., 1965, стр. 126—138. — В том же выпуске (стр. 46—47) приведены критические замечания Хомского. Типологическое значение гипотезы Ингве обсуждается в «Моделях языка». Интересный подход к гипотезе Ингве см.: Ю. А. Шрейдер. Характеристики сложности структуры текста. «Научно-техническая информация», № 11, 1966.

## § 51. НЕКОТОРЫЕ ЗАМЕЧАНИЯ О ТРАНСФОРМАЦИОННОЙ МОДЕЛИ

Идея трансформационного порождения, выдвинутая в теории Хомского и развиваемая целым рядом авторов, основана на следующей предпосылке. Считается, что имеется некоторый подязык, содержащий предложения простой структуры и называемый ядром языка, причем соответствующие фразы называются ядерными, и эти ядерные фразы порождаются грамматикой непосредственно составляющих (например, бесконтекстной грамматикой), а все остальные фразы получаются путем преобразования полученных ядерных (или же их  $\Pi$ -образов). Эту идею можно отразить в следующем определении:

**0.7.10.** Пусть у нас имеется последовательность категориальных пучков  $\Pi_1 \Pi_2 \dots \Pi_n$  (I), полученная на каком-то этапе порождения одного или нескольких ядерных предложений  $g_1, \dots, g_k$ , являющихся реальными фразами, и пусть цепочка, соответствующая этой последовательности, имеет вид  $f = x_1, \dots, x_n$ .

Трансформационным называется правило, сопоставляющее этой последовательности категориальных пучков последовательность категориальных пучков  $\Pi'_1 \Pi'_2, \dots, \Pi'_k$  (II), для которой задана схема составляющих  $\Sigma'$ . Трансформационное правило называется регулярным, если имеется реальная фраза  $f' = y_1, \dots, y_l$ ,  $\Pi$ -образом которой является последовательность (II) и которая  $\Gamma$ -связана с фразой  $f$ .

В данной теории множество фраз формируется не только определением 3.16 (или 3.6), но и еще одним правилом:

**0.7.11.** Если  $(\Sigma, \Pi_1 \Pi_2, \dots, \Pi_n) \rightarrow (\Sigma', \Pi'_1 \Pi'_2, \dots, \Pi'_k)$  есть трансформационное правило, где  $\Sigma$  и  $\Sigma'$  суть схемы составляющих, то любая цепочка  $f$ , имеющая  $\Pi$ -образ, который совпадает с последовательностью категориальных пучков в правой части правила, является фразой в  $L$  и имеет схему составляющих  $\Sigma'$ .

Если употребляемое трансформационное правило — регулярное, то тогда каждая фраза  $f$ , полученная путем применения такого правила, находится в отношении трансформируемости к некоторой последовательности ядерных фраз  $g$ .

Так можно проверить, что фразы *плотник строит дом* и *дом строится плотником* находятся в отношении трансформируемости в смысле определения 5.26. Изложенная же здесь концепция предполагает, что сначала методом непосредственно составляющих производится фраза *плотник строит дом* и лишь потом путем применения трансформационного правила получается фраза *дом строится плотником*. Аналогично во всех других случаях, например при субстантивации (*Колумб открыл Америку*  $\rightarrow$  *открытие Колумбом Америки*), адъективации (*дом сестры*  $\rightarrow$  *сестрин дом*), вербализации (*алый парус*  $\rightarrow$  *парус алел*) и т. п.



Такая интерпретация при всей ее логической привлекательности представляется все же спорной. Преобразования, подобные приведенным заменам действительного залога страдательным, или субстантивации, действительно учат производить каждого школьника, но все они относятся скорее к сознательной деятельности, сознательной работе над языком. Трудно предположить, чтобы ребенок, обучаясь языку<sup>10</sup>, научился естественным способом подобным правилам. Ведь все подобные фразы могут быть легко и естественно порождены и методом непосредственно составляющих. Единственный довод в пользу трансформационного порождения таких фраз в их семантическом тождестве с некоторыми фразами более простого строения. Но именно этот чисто семантический довод представляется нам неубедительным, если он не подкреплён дополнительными чисто синтаксическими доказательствами.

Следует отметить, что в других случаях подобные доказательства были найдены. Имеется целый ряд фраз, порождение которых методом непосредственно составляющих связано с внутренними трудностями структурного характера. Сюда относятся многие случаи употребления однородных членов (ср. указанные в § 18 трудности, связанные с установлением в этом случае отношений синтаксического подчинения), в особенности когда в рассмотрение включаются придаточные предложения.

Приведем следующий пример. При порождении предложения *я видел столько городов, сколько тебе и не снилось* (I) методом непосредственно составляющих можно пользоваться системой правил:

- 1) предложение → (1-е л., ед. ч., И) + (прош. вр., неперех., ед. ч.)
- 2) (прош. вр., неперех., ед. ч.) → (прош. вр., перех., ед. ч.) много + (Р, мн. ч.)

3) *много* → *столько, . . . сколько тебе и не снилось*

Но если правило 3) применить к II-образу предложения *этих денег мне много* (II), то получится запрещенная цепочка \**этих денег мне столько, сколько тебе и не снилось*.

Разница между предложениями I и II в том, что слово *много* в них входит в разные составляющие, одна из которых соответствует именной части фразы, а другая — глагольной, причем существенно вхождение этого слова в одну составляющую с дательным субъекта (*мне*). Однако для выяснения этого необходимо, по-видимому, не просто применять правило 3), но учитывать общую схему составляющих, отражающую историю порождения фразы, а это превосходит возможности модели непосредственно составляющих.

---

<sup>10</sup> То, что порождающие правила должны отражать (или объяснить) этапы обучения ребенка языку, неоднократно подчеркивал сам Хомский, см., например: Н. Х о м с к и й. Объяснительные модели в лингвистике. «Математическая логика и ее применения». Перев. с англ. М., 1965, стр. 246—248.

Таким образом, естественнее считать, что все случаи образования одного предложения из нескольких (однородные члены, слитные и сложные предложения) описываются трансформационной грамматикой<sup>11</sup>, причем для этого достаточно регулярных трансформационных правил.

Но нерегулярные трансформационные правила также имеют глубокий лингвистический смысл. Например, можно надеяться, что с их помощью удастся описать те преобразования, которые имеют место в поэтической речи и которые некоторым нерегулярным образом (не пользуясь, например, аналогией) меняют множество фраз (за счет того резервуара незапрещенных цепочек, о котором говорилось в § 19).

## § 52. КОНФИГУРАЦИОННАЯ МОДЕЛЬ

Как неоднократно подчеркивал Хомский, грамматика в его смысле представляет собой некоторую гипотезу о структуре множества фраз. Формулируя соответствующую гипотезу, мы исходим из интуиции, из наших предварительных знаний о структуре соответствующих фраз.

В этой связи приобретают особый интерес модели, позволяющие хотя бы отчасти формализовать нашу интуицию с тем, чтобы облегчить задачу составления грамматик порождения.

Первой моделью такого типа была конфигурационная модель О. С. Кулагиной<sup>12</sup>. В дальнейшем понятие конфигурации было уточнено А. В. Гладким<sup>13</sup>. Мы дадим здесь оба определения конфигурации (как определения Гладкого, так и определение Кулагиной).

0. 7. 12. Конфигурацией 1-го ранга называется цепочка  $X$ , состоящая более чем из одного элемента и обладающая следующим свойством: существует слово  $x_{(1)}$  (называемое результирующим), такое, что для любых цепочек  $C$  и  $D$  цепочки  $CXD$  и  $C x_{(1)} D$  одновременно фразы или одновременно нефразы.

<sup>11</sup> Такая точка зрения представляется лингвистически более оправданной для всех тех языков, в которых (как в славянских) сложные предложения как правило возникают на базе объединения простых. Что же касается тех языков (типа современных тюркских), где сложное предложение возникает на базе распространения простого (см.: А. П. Рифтин. О двух путях развития предложения в аккадском языке. «Советское языкознание», т. III, 1937), то там более естественно применять порождение методом составляющих. Заметим, что указанная теория А. П. Рифтина, восходящая, по свидетельству В. В. Виноградова, к Л. П. Якубинскому (см. вводную статью В. В. Виноградова в кн.: Л. П. Якубинский. История древнерусского языка. М., 1953, стр. 22—23), вполне естественно формализуется в грамматиках порождения.

<sup>12</sup> О. С. Кулагина. Об одном способе определения грамматических понятий на базе теории множеств. «Проблемы кибернетики», вып. 1. М., 1958.

<sup>13</sup> А. В. Гладкий. Конфигурационные характеристики языков. «Проблемы кибернетики», вып. 10. М., 1963, стр. 252—253.



*Очень милая ... милая дама сидит*

*к раз*

*Милая ... милая дама сидит*

*т раз*

Относительно этого множества цепочка *очень милая* есть конфигурация 1-го ранга с результирующим *милая*; цепочка *милая дама* не есть конфигурация 1-го ранга, ибо во фразе *очень милая дама ласкает кошку* нельзя заменить цепочку *милая дама* на слово *дама*. Относительно данного множества фраз цепочка *серую кошку* есть конфигурация 1-го ранга. Если, однако, добавить фразы, содержащие цепочку *очень серую ... серую кошку*, то *серую кошку* уже не будет конфигурацией 1-го ранга. Цепочка *милая дама* есть конфигурация 2-го ранга. Если добавлены и вышеупомянутые фразы, то *серую кошку* также есть конфигурация 2-го ранга. Заметим, что в этом последнем случае заменять *серую кошку* на *кошку* можно и во фразах, содержащих конфигурацию более низкого ранга *очень милая* (по определению Кулагиной, в таких фразах замена не должна производиться). *Ласкает кошку* есть конфигурация 2-го ранга с результирующим *сидит*, если не включать фраз с цепочкой *ласкает кошку сестры* и конфигурация 3-го ранга с тем же результирующим — в противном случае.

Этот пример наглядно показывает зависимость выделения конфигураций от всего множества фраз. Заметим теперь, что согласно определению конфигурацией 2-го ранга с результирующим *дама* будет не только *милая дама*, но и *очень милая ... милая дама* или *милая ... милая дама* все с тем же результирующим *дама*.

В связи с этим вводится следующее определение.

0. 7. 14. Конфигурация ранга  $r$  называется простой, если она не содержит внутри себя конфигураций ранга  $r$ .

Будем теперь в каждой фразе заменять конфигурации на результирующие. В конце концов мы получим фразу, не содержащую никаких конфигураций.

0. 7. 15. Фраза  $f$  называется неприводимой в языке  $L$ , если она не содержит никаких конфигураций этого языка.

А. В. Гладкий ввел следующее понятие.

0. 7. 16. Язык называется конечно характерным, если число простых конфигураций и неприводимых фраз конечно.

Важность этого понятия иллюстрируется следующими теоремами, доказанными А. В. Гладким.

Т. 7. 3. Для всякого конечно характеризуемого языка существует бесконтекстная грамматика, порождающая все фразы данного языка.

Т. 7. 4. Всякий язык, порождаемый автоматной грамматикой без омонимии, конечно характеризуем.

Т. 7. 5. Существует алгоритм, позволяющий по заданной автоматной грамматике без омонимии находить все простые конфигурации и все неприводимые фразы языка, порождаемого этой грамматикой.

Большим достоинством данной модели с точки зрения лингвистической является введение понятия «ранга». Соответствующее понятие было сформулировано еще Есперсеном<sup>14</sup>. Последний, правда, признавал всего три ранга. В структурной лингвистике понятие ранга было использовано, с одной стороны, де Гроотом<sup>15</sup> и, с другой стороны, Харрисом<sup>16</sup>. При этом де Гроота интересовали отношения внутри некоторой группы слов, а Харриса — возможность путем введения соответствующей индексации расширить границы взаимозаменяемости слова и группы слов.

Другой важной особенностью понятия «конфигурация» является то, что оно позволяет для довольно большого класса фраз получить анализ синтаксических связей<sup>17</sup>, достаточно интересный с лингвистической точки зрения.

0. 7. 17. Два слова  $x$  и  $y$  называются синтаксически родственными, если: 1) для любых двух цепочек  $f$  и  $g$  и для любого числа  $r$  цепочка  $fxg$  есть конфигурация ранга  $r$  тогда и только тогда, когда  $fyg$  есть конфигурация ранга  $r$ , и 2) существует хотя бы одна конфигурация вида  $fxg$  с результирующим  $x$ .

Введение понятия родственных слов важно для установления отношений в группах с переходным глаголом. Естественно считать, что конфигурации вида *вижу девочку* имеют результирующие вида *гуляю*, тем самым существенно отличаясь от конфигураций *вижу хорошо*, где результирующей *вижу* совпал с одним из членов конфигурации. Без понятия родственных слов мы не могли бы анализировать связи внутри группы *вижу девочку* методами, развиваемыми ниже<sup>18</sup>. Теперь же мы можем применять

---

<sup>14</sup> См.: О. Есперсен. Философия грамматики. Перев. с англ. М., 1958, гл. VII.

<sup>15</sup> De Groot. Strukturelle Syntaksis. Den Haag, 1949; Он же. Structural linguistics and syntactic laws, «Word», vol. 5, 1949, № 1.

<sup>16</sup> Z. Harris. Methods in structural linguistics. Chicago, 1951, стр. 265—268.

<sup>17</sup> Идея такого анализа была предложена автором в статье «Формальный и семантический анализ синтаксических связей в языке» (см.: сб. «Применение логики в науке и технике». М., 1960, стр. 128—130; см. также: «Модели языка», стр. 128—130). Недавно была опубликована статья А. В. Гладкого, в которой идея анализа синтаксической связи нашла вполне формализованное выражение (см.: А. В. Г л а д к о й. Об одном способе формализации понятия синтаксической связи. «Проблемы кибернетики», вып. 11. М., 1964, стр. 199—213). Мы воспользуемся здесь некоторыми определениями и формулировками А. В. Гладкого.

<sup>18</sup> Характерно, что в «Моделях языка», где не использовалось понятие родственных слов, группы вида *одержали победу* трактовались как единое целое в схеме зависимостей.

одну и ту же систему определений к конфигурациям вида *очень красивой* или *красивый мост*, с одной стороны, и *вижу девочку*, с другой стороны.

0. 7. 18. Конфигурация  $X$  называется *разложимой*, если для какого-нибудь ее результирующего  $z$  существует представление  $X$  в виде  $X = X_1 y X_2$ , где  $y$  — слово, родственное  $z$ . Слово  $z$  называется в этом случае *главным результирующим*, а слово  $y$  — *потенциальным ядром*.

0. 7. 19. Разложимая конфигурация называется *регулярной*, если в ней только одно потенциальное ядро. В этом случае мы называем часть конфигурации, соответствующую потенциальному ядру, просто *ядром*, а оставшуюся часть — *атрибутом*.

Примеры регулярных конфигураций: *очень маленькая*, где *маленькая* — ядро, а *очень* — атрибут, *маленькая девочка*, где *девочка* — ядро, *маленькая* — атрибут, *вижу девочку* (с главным результирующим *гуляю*), где *вижу* — ядро и *девочку* — атрибут. Таким образом, регулярные конфигурации соответствуют эндосентрическим группам.

Примеры разложимых нерегулярных конфигураций: *маленькая красивая*, где как *маленькая*, так и *красивая* являются потенциальным ядром, *пришел, увидел, победил*, где потенциальным ядром конфигурации является любое из входящих в нее слов.

Примеры неразложимых конфигураций: *за столом* с результирующим *там*, франц. *le mort* с результирующим *il*, чешское *jsem zpíval* с результирующим *zpívali*. Все экзоцентрические группы представляют собой неразложимые конфигурации.

Если бы вся фраза состояла из двучленных регулярных конфигураций, то систему синтаксических подчинений было бы нетрудно получить на основании знания конфигураций. А именно: если каждая конфигурация регулярна и состоит ровно из двух слов, одно из которых атрибут, а другое — ядро, то естественно считать, что каждый атрибут непосредственно зависит от своего ядра.

Практически самым большим препятствием на пути такого анализа являются неразложимые конфигурации, ибо в остальном в пределах простого предложения мы как раз наблюдаем картину, близкую к описанной выше.

В связи с этим полезным оказывается понятие квазиконфигурации, также введенное в цитированной статье А. В. Гладкого.

0. 7. 20. Цепочка  $X$  называется *квазиконфигурацией* ранга  $r$ , если ее можно представить в виде  $X = X_0 Y_1 X_1 Y_2 X_2$ , так что:

- 1) длина цепочки  $Y = Y_1 Y_2$  больше 1;
- 2) хотя бы одна из цепочек  $X_0$ ,  $X_1$  или  $X_2$  непуста;
- 3) каковы бы ни были цепочки  $Z_1$  и  $Z_2$  имеет место:  
а) если  $Z_1 Y Z_2$  фраза, то фраза и  $Z_1 X Z_2$  и

б) если  $Z_1XZ_2$  фраза и не содержит вхождений разложимых конфигураций рангов, меньших  $r$ , пересекающихся с выделенным вхождением  $X$ , но не содержащихся целиком в этом вхождении, то и  $Z_1YZ_2$  фраза.

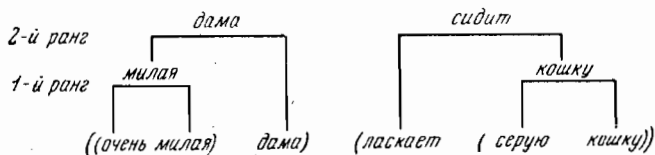
0. 7. 21. Цепочку  $Y$  будем называть ядром квазиконфигурации, а цепочку  $X_0X_1X_2$  — а т р и б у т о м  $X$ .

Теперь мы можем анализировать и фразы, содержащие неразложимые конфигурации. Возьмем фразу *он поднялся на высокий холм*. Здесь на *высокий холм* есть квазиконфигурация с ядром *на холм* и атрибутом *высокий*. Если мы опять будем устанавливать схему зависимостей по принципу, что каждый атрибут синтаксически зависит от своего ядра, то мы получим, что слово *высокий* зависит от группы *на холм*. Иначе говоря, на синтаксическом уровне предложное сочетание должно трактоваться как единый элемент. Мы еще встретимся в ряде других мест с соображениями, требующими именно такого рассмотрения.

Модель конфигурационного анализа обладает, однако, некоторыми существенными недостатками, которые нельзя устранить введением более дифференцированной системы определений.

Эти недостатки были подробно проанализированы Е. В. Падучевой<sup>19</sup> и, в частности, состоят в следующем.

Мы хотели бы, чтобы конфигурационный анализ дал нам систему составляющих некоторой фразы, т. е. дал бы нам возможность восстановить способ ее порождения. Приводившиеся до сих пор примеры, казалось бы, подтверждают, что это возможно. Действительно, конфигурационный анализ фразы *очень милая дама ласкает серую кошку* приводит к результату, соответствующему способу порождения этой фразы.



В связи с этим Е. В. Падучева ввела следующее определение.

0. 7. 22. Язык, в котором любая конфигурация любой фразы является составляющей, называется к о н ф и г у р а ц и о н н ы м.

Используя очень простой метод доказательства, Е. В. Падучева показала, что русский язык не является конфигурационным языком, вернее, что при некоторых вполне оправдываемых требованиях к множеству правильных фраз русского языка это множество не будет конфигурационным. Предположим, что соответствующее множество фраз строится по принципу: в каждой фразе есть подлежащее и в каждой фразе с переходным глаголом есть допол-

<sup>19</sup> Е. В. Падучева. О понятии конфигурации. — ВЯ, 1965, № 1, стр. 56—68.

нение, т. е., например, цепочки *ее не любит* и *лошадь не любит* запрещены. Тогда во фразах *Иван лошадь Петра не любит* и *лошадь Петра не любит* со следующими системами составляющих:

I. (*Иван (лошадь Петра) не любит*)

II. (*лошадь (Петра не любит)*)

конфигурации не совпадают с составляющими. В самом деле, если бы цепочка *лошадь Петра* была бы конфигурацией 1-го ранга, например с результирующим *лошадь* или *ее*, то было бы нарушено условие, выдвинутое выше. Эта цепочка может быть конфигурацией более высокого ранга при условии, что цепочки *Иван лошадь Петра* или *Петра любит* суть конфигурации более низкого ранга. Но первая цепочка не может быть конфигурацией, так как по условию при *любит* должны стоять и субъект и дополнение. Допустим, что *Петра любит* есть конфигурация. Но нельзя заменить *Петра любит* ни на непереходный глагол, например *спит*, ибо фраза I превратилась бы в неправильную, ни на переходный, ибо тогда неправильной стала бы фраза II. Если же заменить *Петра любит* на глагол, который может быть и переходным и непереходным, например *рисует*, то такая замена возможна во фразах I и II, но невозможна, например, во фразе

III. *Он (рисует свою дочь)*,

где *рисует* нельзя заменить на *Петра любит*. Аналогично показывается, что *Петра любит* не может быть конфигурацией более высокого ранга, чем *лошадь Петра*.

Этот пример показателен не только как чисто лингвистический способ доказательства теорем. Он наглядно показывает, что конфигурационный анализ некоторой фиксированной фразы зависит от всего множества допускаемых фраз и меняется в зависимости от него, в то время как система составляющих данной фразы не зависит от систем составляющих других фраз (т. е. от того, какие еще фразы порождаются данной грамматикой).

В § 55 мы рассмотрим модель, в которой анализ фиксированной фразы гораздо меньше зависит от анализа других фраз.

### § 53. СИСТЕМА СОСТАВЛЯЮЩИХ И ДЕРЕВО СИНТАКСИЧЕСКОГО ПОДЧИНЕНИЯ

В § 18 нами обсуждались два отношения синтаксической зависимости, а именно отношение  $D$  и его транзитивное замыкание  $D^+$ , которые отражают систему синтаксических связей во фразе.

В связи с этим интересно исследовать, в какой мере связаны между собой два разных способа представления синтаксической структуры фразы: через систему составляющих и через дерево синтаксического подчинения.

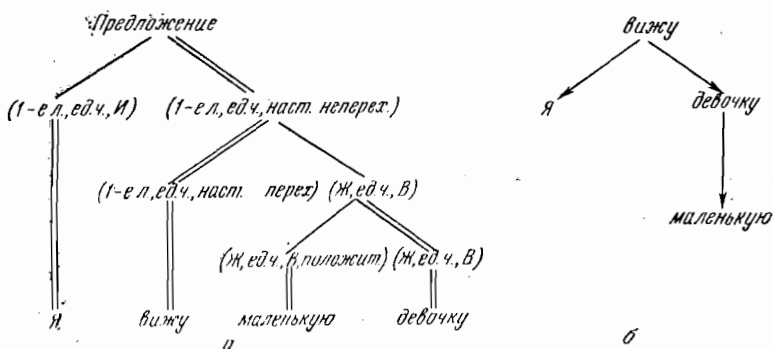


Этот вопрос был подробно рассмотрен Е. В. Падучевой<sup>20</sup>.

Она вводит следующие понятия.

0. 7. 23. Кустом вхождения  $x$  во фразу  $f$  называется поддерево синтаксического подчинения, состоящее из вхождения  $x$  и всех слов  $y$ , таких, что  $x D^+ y$ . Куст вхождения  $x$  называется тривиальным, если в него входит  $y$  такое, что  $y \neq x$ .

Для того, чтобы однозначно переходить от системы составляющих к дереву (или в общем случае — графу) непосредственного синтаксического подчинения, нужно иметь сведения о том, где в каждой составляющей главный, а где подчиненный компоненты. Так, можно договориться, чтобы к главному компоненту вела двойная линия на схеме. Это и сделано ниже в схеме.



Вводится понятие главного слова составляющей.

0. 7. 24. Если составляющая состоит более чем из одного слова, то ее главным словом является то, до которого можно дойти от узла, соответствующего данной составляющей, если спускаться вниз только по двойным линиям.

Правило, сформулированное Падучевой, состоит в следующем: чтобы получить дерево синтаксического подчинения из системы составляющих, достаточно в каждой составляющей провести стрелку от главного слова этой составляющей к главному слову ее подчиненного компонента.

Переход от дерева синтаксического подчинения к системе составляющих производится так. Идя по предложению, мы находим отрезок, соответствующий некоторому минимальному нетривиальному кусту дерева. Закljučаем этот отрезок в скобки и, рассматривая лишь вершину этого куста, вновь повторяем операцию, пока не будут исчерпаны все кусты. Например, наша фраза будет иметь вид: (я вижу (маленькую девочку)).

Здесь сразу же виден один недостаток данного метода, состоящий в том, что сколько бы ветвей ни было у куста, ему все равно

<sup>20</sup> Е. В. Падучева. О способах представления синтаксической структуры предложения. — ВЯ, 1964, № 2, стр. 99.

будет соответствовать ровно одна составляющая. Мы приписывали, однако, данной фразе (стр. 195) более дифференцированную систему составляющих: (я (*вижу (маленькую девочку)*)). Для полного совпадения получающейся системы составляющих с истинной необходимо уметь формировать составляющие в пределах одного куста. Это можно достигнуть, выделив в отношении синтаксического подчинения отдельные подтипы, например предикативную, комплетивную связь и т. п. Но это все равно делается при решении целого ряда задач<sup>21</sup>, поэтому можно считать, что вместе с деревом синтаксического подчинения заданы индексы ветвей, отражающие степень близости к подчиняющему слову. Как показывает Е. В. Падучева, это обеспечивает переход от дерева синтаксического подчинения к системе составляющих.

#### § 54. ОБ ОДНОМ СПОСОБЕ ВОССТАНОВЛЕНИЯ КАТЕГОРИИ НА БАЗЕ РЕАЛЬНЫХ И ЗАПРЕЩЕННЫХ ЦЕПОЧЕК

Конфигурационная модель важна как способ прояснения лингвистической интуиции, лежащей, по-видимому, в основе анализа по составляющим, однако она очень громоздка для практического анализа даже достаточно просто устроенных фрагментов языка (например, фрагментов, конфигурационных в смысле Падучевой). Поэтому целесообразно рассмотреть некоторую более простую модель, сохраняющую часть идей конфигурационного анализа, но применяемую, во-первых, к конечному и достаточно обозримому множеству фраз и, во-вторых, использующую более простые процедуры.

Это, в свою очередь, позволит поставить новые вопросы, в частности, вопрос о возможности вывести категории из рассмотрения фраз. Эта модель возвращает нас к исходным понятиям множества реальных фраз и множества запрещенных цепочек. Мы будем также считать, что реальные фразы сохраняют свойство конфигурационности, в смысле Падучевой. Кроме того, нам задано разбиение на парадигмы. Для наглядности сразу приведем пример.

Реальные фразы:

1. *Мальчик стоит*
2. *Девочка стоит*
3. *Стол стоит*
4. *Мальчик стоит там*

<sup>21</sup> С той точки зрения интересна классификация типов синтаксических отношений, предложенная И. А. Мельчуком для решения задач машинного перевода; см.: И. А. М е л ь ч у к. Об алгоритме синтаксического анализа языковых текстов (общие принципы и некоторые итоги). «Машинный перевод и прикладная лингвистика», № 7, 1962, стр. 48—87; Ол же. Автоматический синтаксический анализ. Новосибирск, 1964, стр. 22—26.

5. Мальчик ругает девочку
6. Мальчик ругает мальчика
7. Мальчик ругает мальчиков
8. Мальчик ругает девочек
9. Мальчик ругает стол
10. Маленький мальчик стоит
11. Мальчик стоит около девочек
12. Мальчик стоит около девочки
13. Девочка стоит около мальчика
14. Очень маленькая девочка стоит
15. Очень маленький мальчик стоит
16. Мальчик стоит около маленькой девочки
17. Мальчик стоит около красивой девочки
18. Девочка стоит около маленького мальчика
19. Маленький мальчик ругает маленькую девочку
20. Мальчик стоит около маленьких девочек
21. Маленький мальчик ругает маленьких девочек

Запрещенные цепочки:

1. \*Девочка стоит около девочку
2. \*Маленький девочка стоит
3. \*Маленькая мальчик стоит
4. \*Мальчик ругает маленькую мальчика
5. \*Мальчик ругает маленькую девочек
6. \*Мальчик ругает маленького мальчиков
7. \*Мальчик стоит около маленькой девочек
8. \*Мальчика стоит
9. \*Мальчик ругает мальчик

Разбиение на парадигмы

$\Gamma_1$	$\Gamma_2$	$\Gamma_3$	$\Gamma_4$
мальчик	девочка	стол	стоит
мальчика	девочки		
мальчиков	девочек		
$\Gamma_5$	$\Gamma_6$	$\Gamma_7$	$\Gamma_8$
ругает	там	маленький	очень
		маленькая	
		маленького	
		маленькой	
		маленьких	

По определению 3.5, выделим все сегменты. Введем теперь понятия ранга фразы и ранга сегмента.

**0. 7. 25.** Назовем реальные фразы минимальной длины фразами нулевого ранга. Присоединим к фразам 0-го ранга все реальные фразы, содержащие ровно на одно слово больше, и выделим на получившемся множестве все сегменты. Назовем их сегментами 1-го ранга. Назовем все фразы, не содержащие других сегментов, кроме сегментов 1-го ранга, фразами 1-го ранга.

Пусть теперь для всех  $i$  ( $i \leq r$ ) определено понятие фразы  $i$ -го ранга и сегмента  $i$ -го ранга.

Возьмем все фразы  $r$ -го ранга и присоединим к ним все реальные фразы, содержащие ровно на одно слово больше, и выделим на получившемся множестве все сегменты. Назовем их сегментами  $(r+1)$ -го ранга. Назовем все фразы, не содержащие других сегментов, кроме сегментов  $i$ -го ранга ( $i=1, 2, \dots, r+1$ ), фразами  $(r+1)$ -го ранга.

**0. 7. 26.** Будем говорить, что слово  $x$  дистрибутивно близко со словом  $y$  по рангу  $r$ , если на фразах ранга  $r$  имеет место отношение  $xQy$  (где  $Q$  — транзитивное замыкание отношения  $P$  в смысле определения 3.4).

Поскольку отношение дистрибутивной близости для любого ранга  $i$  есть отношение типа эквивалентности, то каждое такое отношение  $Q^{(i)}$  разбивает все множество слов на непересекающиеся классы, которые мы будем обозначать соответственно  $Q_1^{(i)}, \dots, Q_n^{(i)}$ .

В нашем примере фразами 0-го ранга будут три первые фразы, а дистрибутивно близки по 0-у рангу и, стало быть, входят в один класс  $Q_1^{(0)}$  слова *мальчик*, *девочка*, *стол*.

Сегментами 1-го ранга в нашем примере будут: *стоит там*, *ругает мальчика*, *ругает девочку*, *ругает девочек*, *ругает стол* — все они развертывают слово *стоит* — и *маленький мальчик* — развертывает слово *мальчик*, *маленькая девочка* — развертывает слово *девочка*.

На множестве фраз 1-го ранга выделяются, например, такие классы дистрибутивно близких словоформ:

$$Q_1^1 = \{\text{стол}\}$$

$$Q_2^1 = \{\text{мальчик}\}$$

$$Q_3^1 = \{\text{девочка}\}$$

Слова *стол* и *мальчик*, входившие относительно фраз 0-го ранга в один класс, входят здесь в два разных класса, так как, например, фраза 1-го ранга *мальчик ругает стол* есть реальная фраза, а цепочка \**мальчик ругает мальчик* запрещена. Слова *мальчик* и *девочка* относятся к разным классам, так как фраза 1-го ранга *маленький*

*мальчик стоит* есть реальная фраза, а цепочка \* *маленький девочка стоит* запрещена.

Интерес представляет новый класс дистрибутивно близких слов, появляющихся на фразах 1-го ранга, а именно:

$$Q_4^1 = \{\text{девочку, девочек, мальчика, мальчиков}\}$$

Сегментами 2-го ранга в нашем примере являются: *около девочки, около мальчика, около девочек, около мальчиков* — все развертывают слово *там* и *очень маленький* — развертывает слово *маленький*. Относительно фраз 2-го ранга выделяются, например, следующие классы:

$$Q_1^2 = \{\text{девочку}\}$$

$$Q_2^2 = \{\text{девочки, девочек, мальчика, мальчиков}\}$$

Слова *девочку* и *девочек* теперь попадают в разные классы, потому что, например, *мальчик стоит около девочек* — реальная фраза 2-го ранга, цепочка \* *мальчик стоит около девочку* — запрещена.

Наконец, в нашем примере имеются следующие сегменты 3-го ранга: *маленькую девочку* развертывает слово *девочку, маленьких девочек* — слово *девочек, маленького мальчика* — слово *мальчика, маленькой девочки* — слово *девочки* и *маленьких мальчиков* — слово *мальчиков*.

В нашем примере нет сегментов ранга более высокого, чем 3-й, и поэтому никакая фраза не может иметь более высокого ранга. Мы можем поэтому рассматривать все реальные фразы как фразы 3-го ранга.

Вообще модель устроена так, что отношение дистрибутивной близости для фраз самого высокого ранга совпадает с отношением дистрибутивной близости для всех фраз, как оно задается определением 3.4.

Теперь мы можем вернуться к нашей основной задаче, а именно, приписать каждому слову две категории (одну парадигматическую, которая, по-видимому, соответствует абстрактной категории рода, а другую, вообще говоря, непарадигматическую, которая, по-видимому, соответствует абстрактному падежу).

Приписывание категорий может осуществляться следующим алгоритмом: мы исходим из предпосылки, что для каждого слова  $x$  и для каждого ранга  $i$  определены все классы дистрибутивной близости  $Q^{(0)}(x), Q^{(1)}(x), \dots, Q^{(n)}(x)$ , где  $n$  — максимальный ранг.

1. Приписать каждому классу нулевого ранга (т. е. всем словам данного класса) свою парадигматическую категорию и перейти к 2.

2. Если некоторому слову  $x$  из некоторого класса  $i$ -го ранга приписана парадигматическая категория  $\kappa$ , то всем классам  $i$ -го

ранга, содержащим хотя бы одно слово из  $\Gamma(x)$ , приписать категорию  $\kappa$  и перейти к 3.

3. Взять очередной класс очередного  $i$ -го ранга. Если этому классу  $Q^{(i)}$  не приспана никакая парадигматическая категория или же  $Q^{(i)}$  есть собственное подмножество класса  $Q^{(i-1)}$ , причем никаким двум классам  $Q_k^{(i)}$  и  $Q_l^{(i)}$ , входящим в  $Q^{(i-1)}$ , не приспаны по пункту 2 разные категории  $\kappa_1$  и  $\kappa_2$ , то приписать  $Q^{(i)}$  новую парадигматическую категорию и перейти к 2. В противном случае оставить старую парадигматическую категорию и перейти к 4.

4. Посмотреть, не исчерпаны ли все классы: если да, то перейти к 5, если нет, то вернуться к 3.

5. Приписать каждому классу 0-го ранга свою синтагматическую категорию и перейти к 6.

6. Взять очередной класс очередного ранга; если данный класс при переходе к  $(i+1)$ -му рангу разбивается на несколько классов, причем каждому из этих классов приспана своя парадигматическая категория, то оставить классам  $(i+1)$ -го ранга синтагматические категории классов  $i$ -го ранга и перейти к 7, в противном случае заменить синтагматическую категорию каждого из классов  $(i+1)$ -го ранга на новую категорию, отличную от категорий других классов, и перейти к 7.

7. Посмотреть, не исчерпаны ли все классы; если нет, то перейти к 6, если да, то остановиться.

Содержательно идея этого алгоритма состоит в следующем: при расширении множества реальных фраз появляются новые слова, синтаксически связанные с данным и поэтому ограничивающие возможности его употреблений. Эти ограничения могут быть связаны с выражением особого согласования, т. е. отграничивать друг от друга отдельные роды (ср. идею «согласовательных классов»). Если же эти ограничения не связаны с парадигматическими категориями, то естественно предположить, что они выражают какие-то синтагматические категории. Это отражено и в структуре алгоритма<sup>22</sup>.

Наш алгоритм неявно предполагает следующее свойство естественного языка:

**0.7.27.** Назовем языком с равномерно суживающимися классами язык, в котором для любого  $i$  имеет место включение  $Q^{(i-1)} \subseteq Q^{(i)}$ .

По-видимому, естественные языки обладают таким свойством, во всяком случае в нашем примере оно выполняется. Дело в том, что практически множество фраз, порожденное множеством фраз  $i$ -го ранга, есть правильный фрагмент и, кроме того, для любого  $i$  класс дистрибутивной близости практически совпадает с классом

<sup>22</sup> Более простое изложение этих правил можно найти в работе: И. И. Ревзин. Лингвистическое моделирование и обучение языку. «Русский язык в национальной школе», 1964, № 6, стр. 61.

Е-эквивалентности на данном фрагменте. Было бы интересно выяснить, при каких абстрактных условиях это свойство имеет место.

Вернемся теперь к нашему примеру.

Классу 0-го ранга, содержащему слова *девочка, мальчик, стол*, будет приписана одна парадигматическая категория (мы обозначим ее через  $\kappa_1$  и будем интерпретировать как «имя существительное») и одна непарадигматическая категория (мы обозначим ее через  $\sigma_1$  и будем интерпретировать как «именительный падеж»). При переходе к 1-му рангу этот класс разбивается на три, причем никакие два из них не получили еще разных категорий, поэтому каждому из новых классов мы припишем новые парадигматические категории  $\kappa_2$ ,  $\kappa_3$  и  $\kappa_4$  (которые соответственно можно интерпретировать как «женский род», «мужской одушевленный род» и «мужской неодушевленный род»), но все эти классы сохраняют категорию  $\sigma_1$ .

Словам *девочку, мальчика, мальчиков* не будет приписано новых парадигматических категорий, зато они получают новую непарадигматическую категорию  $\sigma_2$ , которую можно интерпретировать как винительный падеж. При переходе ко 2-му рангу классы  $Q^{(2)}$  (*мальчик*) и  $Q^{(2)}$  (*девочка*) сохраняются в том же объеме, а значит, сохраняют те же парадигматические категории. Поэтому новые классы:  $Q^{(2)}$  {*девочку*} и  $Q^{(2)}$  {*девочки, девочек, мальчика, мальчиков*} сохраняют те же парадигматические категории и получают новые непарадигматические  $\sigma_3$  (для *девочку* — эту категорию можно интерпретировать как чистый винительный) и  $\sigma_4$  (для *девочки, девочек, мальчика, мальчиков* — эту категорию можно интерпретировать как родительный падеж).

## § 55. ВОССТАНОВЛЕНИЕ ОТНОШЕНИЙ СИНТАКСИЧЕСКОГО ПОДЧИНЕНИЯ. СИНТАКСИЧЕСКИЕ ГРУППЫ

Если конфигурационная модель должна была восстановить истинную систему составляющих путем анализа данной совокупности фраз, то модель, разбираемая в данном параграфе, служит для восстановления истинной схемы синтаксического подчинения для данного фиксированного предложения.

Эта модель предложена чешским математиком Л. Небеским<sup>23</sup>.

Наиболее существенное отличие этой модели от тех, которые до сих пор использовались в синтаксисе, состоит в том, что она явным образом использует понятие смысла предложения. В одной из предыдущих работ<sup>24</sup> автор пытался дать такое видоизменение

<sup>23</sup> L. Nebeský. O jedné formalizaci větného rozboru. «Slovo a slovesnost», XXIII, № 2, 1962, стр. 104—107 (русск. перев.: Л. Небеский. Об одной формализации разбора предложения. «Математическая лингвистика». М., 1964, стр. 145—149).

<sup>24</sup> И. И. Ревзин. Об одной синтаксической модели. — ВЯ, № 2, 1963.

модели Небеского, которое не использует понятия смысла предложения. Поскольку, однако, впоследствии выяснилось, что понятие смысла предложения необходимо и для других целей (ср. § 18), а, с другой стороны, существуют фразы, правильный синтаксический анализ которых можно произвести, только опираясь на смысл предложения, то оказалось более целесообразным явным образом ввести понятие смысла в модель. Это введение можно подкрепить ссылкой и на то обстоятельство, что при составлении грамматики порождения все равно опираются на знание смысла порождаемых фраз.

Переходим к изложению модели Небеского. Пусть нам даны все реальные фразы и на множестве всех реальных фраз задано отношение «смысл фразы  $f$  включается в смысл фразы  $g$ ».

0. 7. 28. Фраза  $f'$  является подфразой  $f$ , если все вхождения (см. определение 3. 2) в  $f'$  присутствуют и в  $f$  и, кроме того, смысл фразы  $f'$  включается в смысл фразы  $f$ .

0. 7. 29. Говорим, что вхождение  $x$  синтаксически подчиняет вхождение  $y$  во фразе  $f$ , если в любой подфразе  $f'$ , в которой присутствует  $y$ , присутствует и  $x$ . Будем говорить, что  $x$  непосредственно подчиняет  $y$ , если  $x$  синтаксически подчиняет  $y$  и не существует  $z$  такого, что: а)  $z \neq x$  и  $z \neq y$  и б) верно, что 1)  $x$  синтаксически подчиняет  $z$ , 2)  $z$  синтаксически подчиняет  $y$  и 3)  $y$  не подчиняет синтаксически  $z$  и  $z$  не подчиняет синтаксически  $x$ .

Примечание. Формулировка пункта б) отличается от формулировки Небеского введением подпункта 3, который у Небеского отсутствует. Это связано со следующим соображением: мы хотим обеспечить для любых  $x$  и  $y$  таких, что  $x$  синтаксически подчиняет  $y$ , наличие последовательности слов  $z_1, \dots, z_n$  таких, что  $z_1 = x$ ,  $z_n = y$  и  $z_i$  непосредственно подчиняет  $z_{i+1}$ . Между тем, если бы мы оставили в б) только пункты 1 и 2, то это не достигалось бы, например, для следующей системы подфраз:  $x$  и  $xuz$ .

Здесь  $x$  синтаксически подчиняет  $y$  и  $x$  синтаксически подчиняет  $z$ , но (если не потребовать условия 3)  $x$  не подчиняет непосредственно ни  $y$ , ни  $z$ .

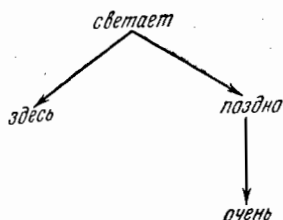
Заметим, что при нашем способе определения отношения непосредственного подчинения граф последнего можно однозначно восстановить из графа отношения синтаксического подчинения путем выбрасывания «лишних» стрелок (т. е. тех, которые пренебрежены условием «б»). Отсутствие пункта 3 в определении Небеского можно объяснить тем, что в случае, когда отношение непосредственного подчинения представляет собой дерево (а в подавляющем большинстве примеров это именно так), это отношение получается из отношения синтаксического подчинения просто путем выбрасывания тех стрелок, которые можно восстановить «по транзитивности» (пункт 3 оказывается ненужным).

Теперь мы сформулируем следующий тезис: отношение непосредственного подчинения, полученное определением 7. 29, совпадает, если оно имеет место между двумя полными словами, с отношением  $D^{(полн)}$ , которое было введено в § 18, а в случае, когда оно имеет место между служебным и полно-



значным словом, оно может быть легко приведено к виду, совпадающему с  $D^{(служ)}$ . Если эта гипотеза верна, то мы получаем возможность для заданной реальной фразы получить отношения  $D = D^{(полн)} \cup D^{(служ)}$  и  $D^+$  и сопоставить каждой фразе граф отношения непосредственного подчинения.

При этом, как мы указали в § 18, мы будем стремиться к тому, чтобы соответствующий граф удовлетворял аксиомам дерева. Так, фраза *здесь очень поздно светает* имеет (кроме себя самой) следующие подфразы: *Здесь поздно светает*, *Здесь светает*, *поздно светает*, *светает*. Поэтому данной фразе соответствует граф:



Заметим теперь, что схема Небеского обладает большей гибкостью, чем схема, предложенная в § 52. Это особенно ярко видно на примере анализа двусоставных предложений. Возьмем фразу *здесь рано наступает ночь*. В зависимости от того, что мы считаем отдельными подфразами этой фразы (а это зависит лишь от того, каким фразам ставится в соответствие законченный смысл), мы получаем разные графы.

Если (как считает сам Небеский, следующий за Шмилауэром<sup>25</sup>, теорию которого он формализует) *наступает* есть подфраза нашей фразы, а *ночь* таковой не является, то граф имеет вид:

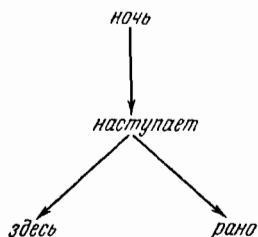


Любопытно, что именно такое дерево (stemma) приписал бы данной фразе Теньер в своей синтаксической теории<sup>26</sup>. Необходимо, однако, отметить, что реально предложения типа *наступает* возможны лишь как эллиптические, а, наоборот, предложения типа *ночь* встречаются (ср. у Блока «Ночь, улица, фонарь, аптека. | Бессмысленный и тусклый свет. | Живи еще хоть четверть века,

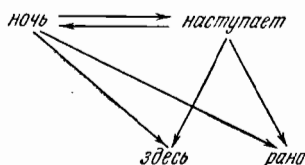
<sup>25</sup> См.: V. Šmilauer. Novočeska skladba. Praha, 1947, стр. 224—225.

<sup>26</sup> См.: L. Tesnière. Éléments de syntaxe structurale. Paris, 1959.

| Все будет так. Исхода нет»). Если принять, что *ночь* есть подфраза нашей фразы, а *наступает* не является ею, то граф будет иметь вид:



Если же считать, что ни *ночь*, ни *наступает* не являются подфразами нашей фразы, что соответствует наиболее типичным ситуациям, то получим граф почти такого же вида, какой предложен в «Моделях языка»:



Четвертый случай, когда и *ночь* и *наступает* суть подфразы нашей фразы, приводит в рассматриваемой модели к графу, противоречащему нашему интуитивному восприятию грамматических связей во фразе. Дело в том, что для лингвистической значимости модели требуется, чтобы граф (во всяком случае фразы, соответствующей простому предложению) был бы связным в смысле определения 2. 20, т. е. для двух разных слов  $x$  и  $y$ , входящих в данную фразу, должно иметь место  $x D^+y$  или  $y D^+x$  или найтись слово  $z$ , подчиняющееся одновременно и  $x$  и  $y$ . Это соответствует интуитивному требованию о наличии единого, не разорванного нигде дерева, соответствующего фразе.

Можно показать, что в рассматриваемой модели имеет место следующее утверждение, устанавливающее критерий связности графа.

**Т. 7. 6.** Для того, чтобы граф фразы  $f$  был связным, необходимо и достаточно, чтобы существовало хотя бы одно слово, присутствующее во всех подфразах фразы  $f$ .

Отсюда следует, что во фразе, граф которой связан, ни одна подфраза не делится на две подфразы (а у нас *ночь наступает* делится на подфразы *ночь* и *наступает*). Таким образом, если ограничиваться только связными графами, то мы должны выбирать для нашей фразы одно из трех первых представлений.

Определить, какой граф лучше соответствует грамматическим фактам, можно только вне модели посредством анализа реального

материала, но самое многообразие выводов в зависимости от разных исходных предпосылок — несомненное достоинство модели. Другое важное преимущество этой модели перед конфигурационной в том, что в ней для анализа некоторой фиксированной фразы не нужно оперировать всем множеством фраз, проверяя распределение некоторой группы слов, как это предлагается конфигурационным методом.

Мы остановимся здесь на графах первого вида, т. е. будем считать, что во множество реальных фраз всегда включены одноэлементные фразы, состоящие из одной личной формы глагола. Это соответствует принципу составления схем синтаксического подчинения в виде дерева, принятому в § 18.

Это означает, что, например, для предложения, разбиравшегося в схеме на стр. 195, мы принимаем следующее множество подфраз:

*я вижу маленькую девочку*  
*вижу маленькую девочку*  
*я вижу девочку*  
*вижу девочку*  
*вижу*  
*я вижу*

Отсюда получаем схему подчинений б) на стр. 208.

До сих пор при определении подфраз можно было не пользоваться понятием смысла, а просто считать, что подфраза получается из фразы путем вычеркивания некоторых слов. Это наводит на мысль построить новую модель, не использующую совсем отношения «смысл фразы  $f$  содержит в себе смысл фразы  $g$ », но зато усиливающую некоторые другие требования. Введем теперь для такой модели аналог понятия подфразы, а именно понятие «упрощения фразы» (здесь в другом смысле, чем в «Моделях языка», где для разбора фиксированной фразы использовался конфигурационный анализ). Определение упрощения мы будем строить индуктивно:

0. 7. 30. а) Фраза  $f$  является упрощением фразы  $f$ ;

б) фраза  $f$ , полученная из упрощения фразы  $f$  путем вычеркивания некоторой цепочки, минимальной в том смысле, что она либо состоит из одного слова, либо вычеркивание более короткой цепочки приводит к цепочке, не являющейся фразой, есть упрощение фразы  $f$ .

С помощью понятия упрощения мы построим теперь новое отношение.

0. 7. 31. Говорим, что вхождение  $x$  доминирует над вхождением  $y$  во фразе  $f$ , если во всех упрощениях фраз  $f$ , в которых присутствует  $y$ , присутствует и  $x$ .

Такой способ восстановления системы синтаксических зависимостей достаточно хорош для многих фраз, для которых отношение доминанции совпадает с отношением синтаксического подчинения. Последнее мы договорились считать идентичным с истинным отношением синтаксического подчинения, введенным в § 18.

В связи с этим можно выделить специальный класс языков, где отношения доминанции и синтаксического подчинения совпадают (аналогично понятию конфигурационного языка, где система конфигураций должна совпадать с истинной системой составляющих).

0. 7. 32. Фраза  $f$  называется доминионной, если в ней отношение доминанции совпадает с отношением синтаксического подчинения. Язык называется доминационным, если каждая его фраза есть доминационная фраза.

В славянских языках фраза, по-видимому, бывает доминационной, когда: а) каждое слово  $x$  своей формой сигнализирует о своей синтаксической зависимости, причем слово, которому подчинено  $x$ , не имеет той же формы, что и  $x$ , и б) во фразе нет омонимичных форм.

Как мы знаем, в подавляющем большинстве случаев в славянских языках свойство а) выполняется. Зато когда данный принцип не соблюдается, мы получаем результаты, противоречащие интуиции. Возьмем, например, фразу: *разберем функции матриц второго порядка*; среди ее упрощений находятся фразы: *разберем функции второго порядка*, *разберем функции матриц*. Этого достаточно, чтобы слова *матриц* и *второго порядка* оказались не связанными отношением доминанции и — что еще хуже — чтобы были установлены некоторые неправильные связи непосредственной доминанции.

Типологически важным является и разбор фраз типа:

*Тучки-кочки переплыли летчики* (В. Маяковский)

Здесь среди упрощений находятся:

*Тучки переплыли летчики*

*Кочки переплыли летчики*

и слова *тучки* и *кочки* оказываются не связанными.

Для славянских языков подобный способ оформления определений достаточно редок, но в целом ряде языков мира определение отличается от определяемого только позицией (ср. конструкции типа *stone wall* в английском языке, определительные конструкции в тюркских языках с препозицией определяемого и в малайско-полинезийских с постпозицией определяемого и т. п.).

Невыполнение свойства б) в славянских языках также часто приводит к тому, что метод доминационного анализа не дает ре-

зультатов. Рассмотрим хотя бы приведенное на стр. 207 предложение: *Иван лошадь Петра не любит.*

Здесь определение 7.29 даст упрощения:

*Иван лошадь Петра не любит*

*Иван Петра не любит*

*Иван лошадь не любит*

*лошадь Петра не любит*

*Иван лошадь не любит*

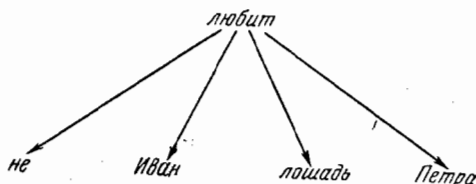
*Петра не любит*

*лошадь не любит*

*не любит*

и т. п.

и следующую бессмысленную схему:

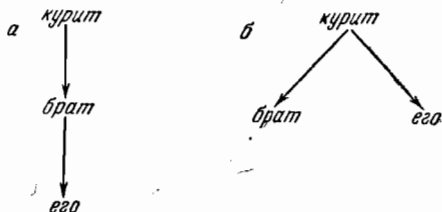


Все такие случаи требуют для адекватного анализа обращения к смыслу.

Если работать только доминационным методом, то нельзя разрешить и случаев омонимии вроде разобранного выше (см. стр. 185, 196). При обращении к смыслу получаем две разные системы подфраз:

- |                             |                           |
|-----------------------------|---------------------------|
| I. 1. <i>брат его курит</i> | II. <i>брат его курит</i> |
| 2. <i>брат курит</i>        | <i>его курит</i>          |
| 3. <i>курит</i>             | <i>брат курит</i>         |
|                             | <i>курит</i>              |

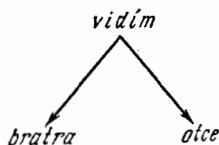
со схемами:



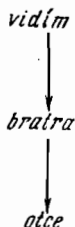
Сравнение метода упрощений и метода подфраз было впервые при несколько другой терминологии праведено Л. Небеским<sup>27</sup>. Рассмотрим один из его примеров. Возьмем фразу *vidím*

<sup>27</sup> Л. Небеский. К одной модели анализа предложения. «The Prague Bulletin of Mathematical Linguistics», 1965, № 2.

bratra otce. Она дает следующие упрощения: *vidím bratra otce*, *vidím bratra*, *vidím otce*, *vidím* и соответственно схему



в то время как истинная схема синтаксического подчинения, отражающая смысл, имеет вид

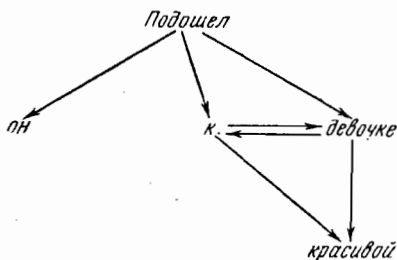


Таким образом, метод доминационного анализа в одних случаях дает адекватные результаты, а в других — нет. Здесь налицо явная аналогия с конфигурационным методом, в одних случаях позволяющим вскрыть структуру составляющих, а в других — нет.

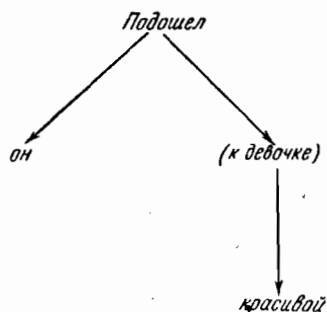
Примеры, приводившиеся в этом параграфе, свидетельствуют, что существует определенная связь между понятиями конфигурационного языка, с одной стороны, и доминационного языка, с другой стороны. Это одно из следствий того, что от системы составляющих можно при некоторых условиях переходить к отношениям синтаксического подчинения и обратно.

Обратимся теперь к одной трудности, возникающей при использовании модели Небеского. Мы уже говорили о том, что нашей целью является восстановление именно отношения *D* из § 18, т. е. мы хотим получить граф, в котором выполняются аксиомы дерева.

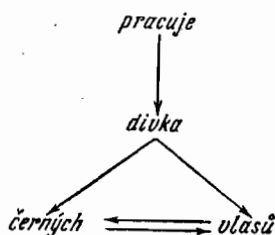
Напомним, что по принятой до сих пор интерпретации слова предлоги являются отдельными словами. Разберем фразу *он подошел к красивой девочке* (ср. стр. 189). Эта фраза дает подфразы: *он подошел к красивой девочке*, *он подошел к девочке*, *он подошел*, *подошел к красивой девочке*, *подошел к девочке*, *подошел*, что приводит к схеме:



Здесь нарушается принцип дерева. Гораздо более естественной была бы схема, если бы слова *к* и *девочке* рассматривались как единый элемент:



Особый интерес с точки зрения этих трудностей модели Небеского представляет анализ вопроса о «неопускаемых определителях»<sup>28</sup>. Предложение *divka černých vlasů pracuje* дает подфразы (по определению 7.28): *divka černých vlasů pracuje*, *divka pracuje*, *pracuje*, что дает схему:

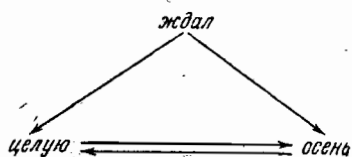


Здесь опять нарушается принцип дерева. Естественной была бы схема:



<sup>28</sup> M. Ivić. Non-omissible Determiners in Slavic Languages. «Proceedings of the ninth International Congress of Linguists». The Haag, 1964, стр. 476.

где *сегнућ* *vlasu* выступает как единый элемент на схеме. Особенно интересен разбор следующей сербской фразы, взятой из работы М. Ивич *секао је село рголесе* (мы для наглядности разберем соответствующий русский пример: *ждал целую осень*). Имеем следующие подфразы (по определению 7.23): *ждал целую осень, ждал*, со схемой:



По определению 7.23, *ждал осень* не есть подфраза данной фразы, так как ее смысл не есть часть смысла фразы *ждал целую осень* (правда, в русском языке для выражения объекта при переходном глаголе *ждать* скорее была бы употреблена форма *ждал осени*, но соответствующий сербский пример, приведенный выше, однозначно воспринимается как сочетание переходного глагола с объектом).

Определение 7.27, не учитывающее смысла, дает здесь лучший результат:



Это не случайность. Имеется, по-видимому, целый ряд случаев, где смысловые связи расходятся с чисто синтаксическими таким образом, что доминационный анализ дает лучшие результаты, чем метод подфраз. Речь идет о фразах типа:

- I. *Он любит читать книги с картинками*
- II. *Он умеет продавать хорошие товары*

Как показала А. Вержбицка<sup>29</sup>, смысл первой из этих фраз существенно отличается от смысла фраз: *он любит читать книги, он любит читать, он любит, любит*.

<sup>29</sup> Устное сообщение в Институте славяноведения в декабре 1965 г.



Поэтому при помощи системы подфраз мы не получим схемы синтаксической зависимости, отвечающей интуиции. С помощью доминационного анализа мы, однако, легко получим схему зависимостей, соответствующую интуиции, ибо все указанные выше фразы являются упрощениями фразы I.

Все же приведенные фразы существенно отличаются от фраз типа *Он подошел к красивой девочке* тем, что хотя бы один из методов (метод подфраз или метод упрощений) приводит для них к схеме дерева, в то время как оба метода дают для последней фразы схему (стр. 221), где, как мы видели, принцип дерева нарушается. В связи с этим введем следующее определение.

**0. 7. 33.** Назовем синтаксической группой фразы  $f$  совокупность вхождений  $x_1, x_2, \dots, x_n$ , принадлежащих  $f$  и таких, что для любых  $x_i$  и  $x_j$  имеет место:

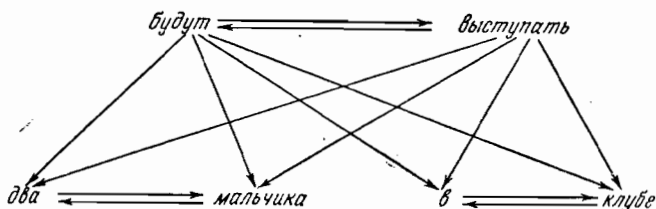
а)  $x_i$  непосредственно подчиняет  $x_j$  и  $x_j$  непосредственно подчиняет  $x_i$ ;

б) для любого  $y$  из  $f$ , если  $y$  непосредственно подчиняет  $x_i$ , то  $y$  непосредственно подчиняет  $x_j$  и наоборот, если  $x_i$  непосредственно подчиняет  $y$ , то  $x_j$  непосредственно подчиняет  $y$ .

**0. 7. 34.** Отобразим множество вхождений на множество синтаксических групп фразы и установим следующее отношение подчинения между синтаксическими группами фразы: группа  $X$  подчиняет группу  $Y$ , если хотя бы одно из слов, входящих в  $X$ , непосредственно подчиняет хотя бы одно из слов, входящих в  $Y$ .

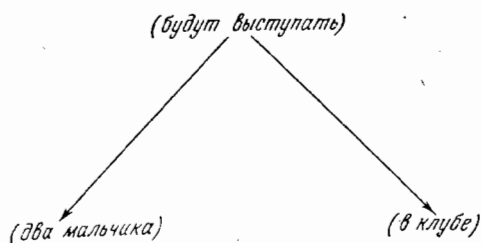
Легко видеть, что это отображение есть гомоморфное отображение, сохраняющее все отношения за пределами групп.

Покажем на примере, как происходит образование синтаксических групп. Возьмем фразу *два мальчика будут выступать в клубе* со следующей системой подфраз: *два мальчика будут выступать в клубе, два мальчика будут выступать, будут выступать в клубе, будут выступать* (здесь мы не будем считать, что *будут* есть подфразы нашей фразы, ибо она не содержит никакого смысла, относящегося к этой фразе). Схема синтаксического подчинения по определению 7. 24 имеет вид:



Образует синтаксические группы в смысле определения 7. 33: *будут выступать, два мальчика, в клубе*. По определению 7. 34, на множестве синтаксических групп имеется следующее отно-

шение синтаксического подчинения:



Сравнение этих схем наглядно показывает гомоморфизм, о котором шла речь, а именно, вторая схема является гомоморфным образом первой схемы. Этот гомоморфизм достигается склеиванием тех вершин, которые связаны с одними и теми же вершинами.

В § 9 мы говорили о том, что гомоморфное отображение отображает лишь существенные связи. Это наглядно проявляется и в данном случае. На гомоморфном образе сложная и запутанная система отношений сразу же проявляется, поскольку все отношения между отдельными членами группы здесь отражены в гораздо более общем виде.

Ниже мы увидим, что имеется ряд чисто лингвистических соображений, заставляющих рассматривать синтаксические отношения не между отдельными вхождениями слов, а между синтаксическими группами.

## § 56. ПОНЯТИЕ ТРАНСПОЗИЦИИ И ПРИНЦИП ПРОЕКТИВНОСТИ

В синтаксических моделях мы всегда имеем дело по крайней мере с двумя синтаксическими отношениями, а именно, с отношением синтаксического подчинения  $D^+$ , а также с отношением линейного порядка  $O$  ( $x$  правее  $y$  во фразе  $f$ ).

Интересно отделить первое отношение от второго и рассматривать некоторые свойства фразы, не связанные с линейным порядком. Мы подойдем к этому при помощи следующего понятия.

0. 7. 35. Фраза  $f'$  находится в отношении транспозиции к фразе  $f$ , если: а) обе они состоят из тех же слов, причем для любого слова  $x$  верно, что если имелось  $k$  вхождений этого слова  $x$  в одну из фраз, то имеется столько же вхождений этого слова в другую фразу и б) обе фразы изоморфны относительно отношения синтаксического подчинения.

Отношение транспозиции рефлексивно, симметрично и транзитивно. Оно разбивает все множество фраз на непересекающиеся подмножества, причем каждое такое подмножество состоит из фраз, различающихся между собой только порядком слов при том же дереве непосредственного синтаксического подчинения.

Каждое такое множество фраз можно рассматривать как одну абстрактную фразу. Именно к этому сводится идея, предложенная одновременно С. К. Шаумяном<sup>30</sup> и Г. Б. Карри<sup>31</sup> и состоящая в различении фенотипического уровня, на котором рассматриваются конкретные фразы, и генотипического уровня, где фраза определяется только своим деревом вне зависимости от реального порядка слов.

Переход от некоторой фразы к находящейся в ней в отношении транспозиции, или транспонированной, всегда может быть представлен как последовательная перестановка двух рядом стоящих слов. При этом, разумеется, может получиться так, что какая-то из промежуточных перестановок делает цепочку неправильной. Поэтому типологическое значение может иметь следующее определение.

**0. 7. 36.** Индексом транспозиции двух фраз  $f$  и  $g$ , находящихся в отношении транспозиции, называется минимальное число не находящихся к ним в отношении транспозиции цепочек, получающихся при переходе от  $f$  к  $g$  путем перестановки на каждом шаге двух рядом стоящих слов.

Возьмем фразы *сегодня я видел очень хорошую книгу* и *я видел очень хорошую книгу сегодня*. Переход от первой ко второй осуществляется следующим образом:

1. *Сегодня я видел очень хорошую книгу*
2. *Я сегодня видел очень хорошую книгу*
3. *Я видел сегодня очень хорошую книгу*
4. *Я видел очень сегодня хорошую книгу*
5. *Я видел очень хорошую сегодня книгу*
6. *Я видел очень хорошую книгу сегодня*

Только одна из этих цепочек (а именно, 4) или не фраза, или же имеет другое дерево и поэтому не находится в отношении транспозиции с нашими фразами. Индекс транспозиции данной пары равен единице. Хотя фразе 5 можно приписать дерево, отличное от дерева фраз 1 и 6 (с отношением *хорошую D сегодня*), одно из деревьев, приписываемых данной фразе, совпадает с деревом фраз 1 и 6. Этого достаточно, чтобы не увеличивать индекс транспозиции. Вообще, по-видимому, очень трудно найти в русском языке простые предложения (т. е. такие, для которых мы разбирали схемы синтаксического подчинения и большинство моделей данной главы), имеющие высокий индекс транспозиции.

<sup>30</sup> См., например: С. К. Шаумян. О логическом базисе лингвистической теории. «Проблемы структурной лингвистики». М., 1963, стр. 3—8.

<sup>31</sup> Н. В. Сургу. Some Logical aspects of Grammatical Structure. — PSAM, vol. 12. «Structure of Language and its Mathematical Aspects», 1961, стр. 56—68 (русск. перев. — «Новое в лингвистике», вып. IV. М., 1965, стр. 97—116).

Другое положение, например, в немецком языке. В частности, возьмем похожие фразы: *In der Bücherei sah ich ein sehr interessantes Buch* и *Ich sah ein sehr interessantes Buch in der Bücherei*.

Если считать, что все транспонированные фразы, начинающиеся с *sah ich*, являются реализацией той же абстрактной фразы, то, кроме них, имеется всего лишь одна промежуточная фраза с тем же деревом (*sah ich in der Bücherei ein sehr interessantes Buch*) и индекс транспозиции равен 17.

Такой показатель с лингвистической точки зрения слишком высок, ибо лингвистически значимо не положение отдельного слова, а положение группы слов. Здесь мы вновь сталкиваемся с необходимостью рассмотрения синтаксических групп. Если же места во фразе в соответствии с лингвистической традицией будут фиксированы не для слов, а для групп, то в русском примере индекс транспозиции будет равным нулю, а в немецком будет равен единице (или двум, если считать, что транспозиция не должна менять интонации). Это различие уже более соответствует интуиции.

Теперь мы вернемся к одновременному рассмотрению двух отношений, а именно отношения линейного порядка (вхождение  $x$  правее  $y$  во фразе  $f$ ) и отношения синтаксического подчинения.

Заметим, что с помощью двухместного отношения « $x$  правее  $y$ » легко формулируется новое трехместное отношение « $x$  находится между  $y$  и  $z$ », а именно это имеет место, когда:

- а)  $y$  правее  $x$ , а  $z$  левее  $x$  или
- б)  $z$  правее  $x$ , а  $y$  левее  $x$ .

Этим отношением «между» мы будем пользоваться при обсуждении принципа «проективности», играющего все более важную роль в моделировании синтаксиса<sup>32</sup>.

**Примечание.** В некоторых работах вместо термина «проективность» употребляется термин «конфигурационность»<sup>33</sup>, однако нам этот термин не годится, ибо понятие «конфигурации» использовано нами в другом смысле.

**0. 7. 37.** Фраза  $f$  удовлетворяет критерию сильной проективности (или является сильно проективной), если для любых трех вхождений  $x$ ,  $y$  и  $z$  из того, что имеет место  $xD^+y$  и кроме того  $z$  находится между  $x$  и  $y$  следует, что имеет место  $xD^+z$ .

**0. 7. 38.** Фраза  $f$  называется слабо проективной, если в ней для любых четырех слов  $x$ ,  $y$ ,  $z$  и  $w$  из того, что имеет место  $xD^+y$  и кроме того  $zDw$ , причем  $z$  находится между  $x$  и  $y$ , следует, что  $w$  также находится между  $x$  и  $y$ .

Слабая проективность сводится к требованию, чтобы стрелки синтаксического подчинения не пересекались.

<sup>32</sup> Ю. А. Шрейдер. Свойства проективности языка. «Научно-техническая информация», 1964, № 8; L. Hirschberg, L. Lynch. Discussions sur l'hypothèse de projectivité. Euratom. — CETIS, № 35, October, 1961.

<sup>33</sup> С. Я. Фиталов. О моделировании синтаксиса в структурной лингвистике. «Проблемы структурной лингвистики». М., 1962, стр. 105—106.

**Примечание.** Наряду с термином «слабая проективность» мы будем пользоваться более общим термином «проективность», особенно в контекстах, не противопоставляющих оба вида проективности.

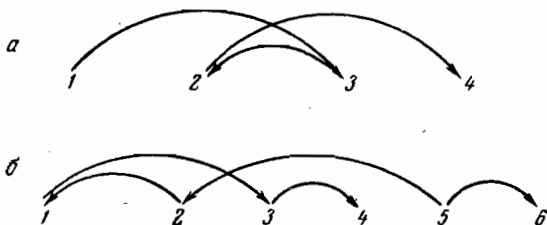
Это свойство интересует нас здесь с той точки зрения, что нарушение его во многих случаях приводит к фразам, отклоняющимся от интуитивно ощущаемых нами норм синтаксиса некоторого конкретного языка, и проективность может рассматриваться как универсалия<sup>34</sup>.

Тем самым можно надеяться получить объективный критерий правильности фразы.

Так, в указанной работе Фитиалова были приведены примеры фраз, порядок слов в которых интуитивно ощущается как неправильный:

а) *обсуждались<sub>1</sub> завершённые<sub>2</sub> работы<sub>3</sub> писателями<sub>4</sub>*

б) *представленные<sub>1</sub> картины<sub>2</sub> на<sub>3</sub> выставку<sub>4</sub> пользовались<sub>5</sub> успехом<sub>6</sub>*



Нарушение проективности может служить в этом случае экспликатом интуитивного понятия неправильности порядка слов.

**Примечание.** Надежда найти в свойстве проективности объективный инструмент определения правильности порядка слов во фразе стала еще более оправданной после исследований Ю. А. Шрейдера по установлению степеней проективности.

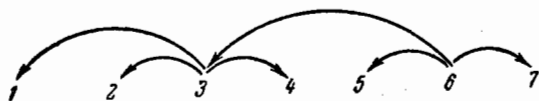
В частности, в исследованиях Шрейдера было показано, что можно вполне естественно распространить принцип проективности на случаи употребления однородных членов. Дело в том, что если не пользоваться соглашением о расстановке стрелок непосредственного подчинения, принятом в § 18, то при естественной расстановке стрелок может стать непроективной вполне обычная фраза:

*эти<sub>1</sub> мальчики<sub>2</sub> и<sub>3</sub> девочки<sub>4</sub> идут<sub>5</sub> и<sub>6</sub> болтают<sub>7</sub>*



<sup>34</sup> Именно так трактует это свойство Лесерф, один из первых исследователей этого вопроса (см.: И. Лесерф. Применение программы и модели конфликтной ситуации к автоматическому синтаксическому анализу естественных языков. «Научно-техническая информация», 1963, № 10, стр. 42—50.

При введенном в § 18 соглашении принцип проективности выполняется.

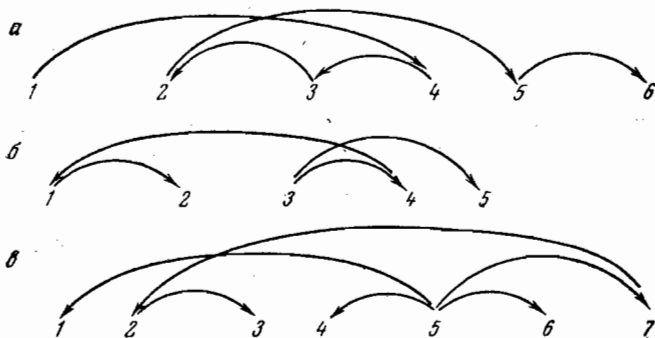


Ощущение неправильности, возникающее при чтении целого ряда непроективных фраз, по-видимому, объясняется тем, что в них нарушены естественные принципы организации фразы таким образом, чтобы связанные по смыслу слова стояли рядом. Эти принципы, несомненно, связаны со способом порождения.

Как было показано в ряде работ <sup>35</sup>, всякая фраза, произведенная грамматикой непосредственно составляющих, должна допускать простую скобочную запись и быть проективной. Весьма знаменательно то обстоятельство, что те непроективные фразы, которые ощущаются как вполне естественные, могут быть объяснены как произведенные грамматикой трансформационной.

Рассмотрим примеры таких фраз, приводимые Л. Н. Иорданской <sup>36</sup> (мы ограничимся только примерами простых предложений, ибо сложные предложения по нашей концепции — заведомо получены трансформационным способом):

- а) Это<sub>1</sub> более<sub>2</sub> понятная<sub>3</sub> книга<sub>4</sub>, чем<sub>5</sub> та<sub>6</sub>
- б) О<sub>1</sub> докладе<sub>2</sub> дала<sub>3</sub> отзыв<sub>4</sub> комиссия<sub>5</sub>
- в) Почему<sub>1</sub> из<sub>2</sub> многих<sub>3</sub> этим<sub>4</sub> занимаются<sub>5</sub> лишь<sub>6</sub> некоторые<sub>7</sub>



Сравнительные обороты естественнее всего представить как возникающие из двух предложений, а остальные предложения имеют явную инверсию (инверсию лучше всего описывать как один из видов трансформаций).

Особенно интересно рассмотрение фактов поэтической речи, для которой инверсия является самой простой трансформацией.

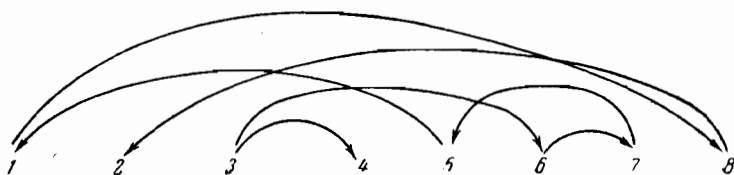
<sup>35</sup> См., например: Ю. А. Шрейдер. Указ. соч.

<sup>36</sup> Л. Н. Иорданская. О некоторых свойствах правильной синтаксической структуры. «Вопросы языкознания», 1963, № 4.

Известно, что в ряде поэтических систем, например в античной поэзии или в польской силлабической поэзии XVII—XVIII веков, в русской поэзии того же времени отклонения от нормального прозаического порядка слов и, в частности, от принципа проективности были очень частым явлением.

Достаточно вспомнить «Метаморфозы» Овидия, где с первой строки непроективность воспринимается как своего рода поэтическая доминанта:

In<sub>1</sub> nova<sub>2</sub> fert<sub>3</sub> animus<sub>4</sub> mutatas<sub>5</sub> dicere<sub>6</sub> formas<sub>7</sub> corpora<sub>8</sub>



Очень характерны непроективные фразы для польского тринадцатисложного силлабического стиха XVIII века, например у поэта С. Трембицкого («Zofiówka»):

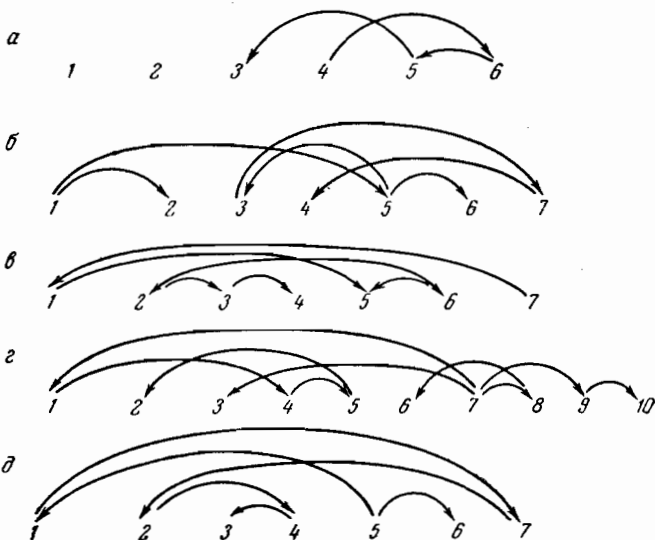
а) Doład<sub>1</sub> jeszcze<sub>2</sub> wieśniacza<sub>3</sub> grunt<sub>4</sub> socha<sub>5</sub> rozjęty<sub>6</sub>

б) wymienia<sub>1</sub> złotem<sub>2</sub> | Uroszony<sub>3</sub> rolniczym<sub>4</sub> owoc<sub>5</sub> ziemi<sub>6</sub> potem<sub>7</sub>

в) A na<sub>1</sub> powinnej<sub>2</sub> dla<sub>3</sub> mnie<sub>4</sub> dodatek<sub>5</sub> ofiary<sub>6</sub> | ... zameñ<sub>7</sub>

г) Między<sub>1</sub> morskiego<sub>2</sub> niegdyś<sub>3</sub> królownamy<sub>4</sub> stanu<sub>5</sub> | Cudnej<sub>6</sub> była<sub>7</sub> urody<sub>8</sub> wnuczka<sub>9</sub> Oceanu<sub>10</sub>

д) W<sub>1</sub> mężów<sub>2</sub> ludzkości<sub>3</sub> pełnych<sub>4</sub> uczu<sub>5</sub> się<sub>6</sub> szeregu<sub>7</sub>



Эти примеры нагромождения непроективности, любезно сообщенные автору проф. М. Р. Майеновой, ясно свидетельствуют о намерении автора построить фразу иначе, чем в прозаическом тексте. Проф. Майенова при этом обратила внимание автора на то, что возможности такой свободы порядка слов обусловлены именно силлабической системой и что при переходе к силлаботоническому стиху, где возрастает роль фразового ударения, непроективные построения уже не встречаются.

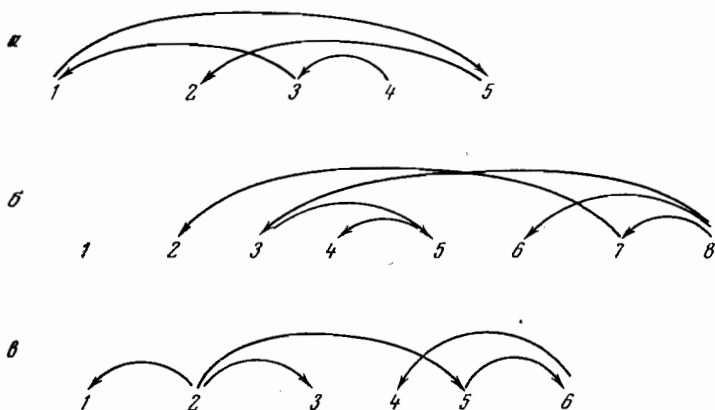
Возможно, однако, что здесь играло роль и то обстоятельство, что более развитая и тонкая система поэтизации текста делала ненужным столь грубое отклонение от того, что воспринималось как норма для прозаического текста.

Нечто подобное наблюдалось и в русской поэзии XVIII века, причем характерно, что непроективность вполне уживается с силлаботонической системой стихосложения, ср.:

а) У<sub>1</sub> Финских<sub>2</sub> спать<sub>3</sub> залег<sub>4</sub> озер<sub>5</sub> (Ломоносов)

б) И<sub>1</sub> всем<sub>2</sub> из<sub>3</sub> своего<sub>4</sub> пера<sub>5</sub> богатства<sub>6</sub> смертным<sub>7</sub> проливашь<sub>8</sub> (Державин)

в) Он<sub>1</sub> тянет<sub>2</sub> руку<sub>3</sub> дам<sub>4</sub> к<sub>5</sub> устам<sub>6</sub> (Он же)



Как ни объяснять приведенные примеры, очевидно, что непроективность в них свидетельствует об их трансформированном характере.

Имеется, однако, ряд случаев, где появление непроективности не может быть объяснено трансформированным характером фраз, где связанные по смыслу слова стоят рядом и фраза воспринимается как вполне естественная.

Возьмем сначала фразу, которую Ю. А. Шрейдер приводит в качестве примера проективной.

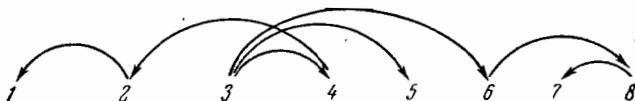


Вечером<sub>1</sub> в<sub>2</sub> новом<sub>3</sub> клубе<sub>4</sub> будет<sub>5</sub> показано<sub>6</sub> эстрадное<sub>7</sub> обозрение<sub>8</sub>



Эта фраза вполне может быть произведена грамматикой непосредственно составляющих. Заметим, однако, что более естественный порядок слов при той же расстановке стрелок приводит к непроективному построению:

Эстрадное<sub>1</sub> представление<sub>2</sub> будет<sub>3</sub> показано<sub>4</sub> вечером<sub>5</sub> в<sub>6</sub> новом<sub>7</sub> клубе<sub>8</sub>



Разумеется, эта фраза может быть произведена грамматикой непосредственно составляющих, но тогда в ней должна быть другая расстановка стрелок, например, если *вечером* и *в* зависят бы не от *будет*, а от *показано*.

Таким образом, оказывается, что при почти тождественной системе составляющих (разница лишь в порядке элементов) приходится допускать две разных расстановки стрелок в зависимости от того, предшествует ли группа сказуемого группе подлежащего или следует за ней.

Представляется, что лучшим выходом из создавшегося тупика было бы признание того факта, что существует по крайней мере два разных уровня рассмотрения текста: а) морфологический и б) синтаксический и что отношения синтаксического подчинения, а вместе с ними и принцип проективности действительно в пределах одного уровня — в данном случае именно синтаксического, для которого, например, слова *будет* и *показано* составляют один элемент и отношения устанавливаются не между словом *будет* (или словом *показано*) и другими словами фразы, а между всей группой *будет показано* и другими словами.

Это разделение уровней можно было бы сопоставить с тем, что ранее было сказано о разнице между автоматной грамматикой и грамматикой непосредственно составляющих.

Как мы отмечали в § 49, морфологический уровень лучше всего моделируется автоматной грамматикой.

Напомним, что автоматная грамматика дает некоторый анализ производимой последовательности, например, разбивает ее на

части (ср. вопрос о функции «пустых слов», разбираемый в § 49), но особенность этого анализа в том, что он говорит лишь о возможности связи данного последующего элемента с предыдущим, не вскрывая отношений синтаксического подчинения. Возможность достаточно хорошо описывать оба типа единиц, а именно слова и словосочетания, этой моделью является формальной характеристикой их, которую интересно сопоставить с их общей семантической ролью: и те и другие выполняют чисто номинативную функцию. Поэтому можно считать, что автоматная грамматика описывает уровень единиц, описывающих действительность в нерасчлененном виде и выступающих единными блоками на следующем уровне — уровне формирования предложения<sup>37</sup>. Этот уровень моделируется грамматикой непосредственно составляющих, на котором каждое высказывание получает явную или неявную субъектно-предикатную структуру (благодаря членению на составляющие) или соответствующую ей систему отношений синтаксического подчинения (в следующем параграфе будет рассмотрен пример грамматики валентностей, которая и даст подобную систему). Именно на этом уровне, по-видимому, нарушение принципа проективности связано с ощущением неправильности.

Наконец, на третьем уровне, трансформационном, могут возникнуть различные инверсии, нарушающие принцип проективности, но необязательно приводящие к ощущению неправильности предложения.

Описанная схема трех уровней возникла из чисто теоретических соображений, связанных с принципом проективности. Разделение морфологического и синтаксического уровней вполне соответствует лингвистической традиции<sup>38</sup>. Что касается второго разграничения, то оно уже неоднократно выдвигалось в связи с другими соображениями, ср. например, различие генотипического и фенотипического уровней в упомянутой гипотезе Шаумяна—Карри<sup>39</sup>. Несколько отличными соображениями руководствовался Д. Уорт<sup>40</sup>, предлагая разграничить соответствующие уровни в связи с анализом славянского языкового материала. Особо следует упомянуть работы современных продолжателей идей пражского лингвистического кружка, в частности Фр. Да-

---

<sup>37</sup> Ср. чрезвычайно интересную теорию Э. Бенвениста об интегративных отношениях, возникающих при переходе от одного уровня к другому: Э. Бенвенист. Уровни лингвистического анализа. Перев. с франц. «Новое в лингвистике», вып. IV. М., 1965, стр. 441—444.

<sup>38</sup> Неразделение этих уровней в американской дескриптивной лингвистике часто критиковал Р. О. Якобсон, который уже давно подчеркивал необходимость разделения этих уровней в структурной лингвистике.

<sup>39</sup> См. работы, указанные в сносках 30 и 31 в начале этого параграфа.

<sup>40</sup> Д. С. Уорт. Об отображении линейных отношений в порождающих моделях языка. — ВЯ, 1964, № 5, стр. 46—58.

неша <sup>41</sup> и П. Сгалла <sup>42</sup>. В последних работах подчеркивается необходимость различать семантический уровень предложения (отражаемый, в частности, системной составляющей) и уровень организации предложения, т. е. трансформационный (актуальное членение предложения, интонационное оформление и т. п.).

Однако во всех указанных работах концепция трех уровней возникает чисто теоретически и не реализована в действующей модели. Имеется, однако, одна работа, в которой эти три уровня естественно возникли в процессе построения действующей модели синтеза русского предложения. Речь идет о замечательной работе, в которой сначала отдельные словосочетания заготавливаются путем заполнения мест в фиксированной таблице (что эквивалентно автоматной грамматике), затем при помощи грамматики валентности (близкой к разбираемой в следующем параграфе) формируются предложения и, наконец, действуют (трансформационные по своей природе) правила расстановки слов в связи с актуальным членением предложения, а именно выделением логического подлежащего <sup>43</sup>.

## § 57. ВАЛЕНТНОСТИ

Одновременное исследование отношений синтаксического подчинения и отношений линейного порядка может быть полезно и для рассмотрения моделей порождения, использующих близкие понятия в качестве исходных (идея построения моделей порождения на основе отношений подчинения предложена в ряде работ <sup>44</sup>). Определим следующие типы отношений.

0.7.40. Мы будем говорить, что слово  $x$  подчиняет справа (слева) слово  $y$  в языке  $L$ , если существует хотя бы одна фраза  $f$  такая, что  $xDu$  в  $f$  и  $x$  расположен правее  $y$  (соответственно  $x$  расположен левее  $y$ ) в  $f$ . Мы будем говорить, что категориальный пучок  $\Pi_1$  подчиняет справа (слева) категориальный пучок  $\Pi_2$ , если существуют  $x \in \Pi_1$  и  $y \in \Pi_2$  такие, что  $x$  подчиняет справа (слева)  $y$ . Категориальный пучок  $\Pi$  называется

---

<sup>41</sup> Фр. Д а н е ш. Опыт теоретической интерпретации синтаксической омонимии. — ВЯ, 1964, № 6, стр. 3—16; Fr. D a n e š. A three-level approach to syntax. «Travaux linguistiques de Prague», 1. Prague, 1964, стр. 225—240.

<sup>42</sup> Р. С г а л л а. Zur Frage der Ebenen im Sprachsystem. «Travaux linguistiques de Prague», 1, 1964, стр. 95—106.

<sup>43</sup> И. А. М е л ь ч у к. Порядок слов при автоматическом синтезе русского текста (предварительное сообщение). «Научно-техническая информация», 1965, № 12, стр. 36—44.

<sup>44</sup> См., например: М. И. Б е л е ц к и й. Модель русского языка, описывающая простые предложения без однородности. «Научно-техническая информация», 1964, № 7, стр. 37—42; D. H a u s. Dependancy Theory a formalism and some observation. «Language», vol. 40, № 4, 1964, стр. 510—514.

абсолютно-подчиняющим, если ни для одного слова  $x \in \Pi$  не найдется такого слова  $y$ , что  $y$  подчиняет  $x$  справа или слева.

**0.7.41.** Назовем валентностью<sup>45</sup> слова  $x$  набор всех категориальных пучков  $\Pi(y)$  и  $\Pi(z)$  таких, что  $\Pi(x)$  подчиняет справа  $\Pi(y)$  и  $\Pi(x)$  подчиняет слева  $\Pi(z)$ .

В качестве примера рассмотрим следующее извлечение из таблицы валентностей.

$\Pi(x)$	Подчиненный категориальный пучок	
	справа	слева
1-е л., ед. ч., наст., перех.	1-е л., ед. ч., И	В, любое число, любой род
ед. ч., В, ЖО	ед. ч., В, Ж, положит.	—
ед. ч., В, МО	ед. ч., В, М, положит.	—
3-е л., ед. ч., наст., перех.	И, ед. ч., любой род	В, любое число, любой род
любой падеж, любое число, любой род и т. п.		Р, любое число, любой род

Пусть такая таблица построена для всего русского языка (в более полном виде такая таблица содержится в работе М. И. Белецкого).

Можно было бы предложить модель порождения, основная идея которой будет изложена ниже (рассматриваются лишь общие принципы модели, которая в некоторых своих чертах была бы похожа на модель Белецкого, хотя и отличалась бы от нее).

Порождающий процесс содержит следующие простые операции.

А  $(x, y)$ : если  $\Pi_1$  и  $\Pi_2$  — выбранные пучки, и  $\Pi_1$  подчиняет  $\Pi_2$  слева (соответственно справа), причем  $x \in \Pi_1$ , а  $y \in \Pi_2$ , то в выписанной цепочке  $x$  заменяется на  $(xy)$  или соответственно на  $(yx)$  и

Б  $(x, y)$ : если  $\Pi_1$  и  $\Pi_2$  — выбранные пучки, и  $\Pi_1$  подчиняет  $\Pi_2$  слева (соответственно справа), причем  $x \in \Pi_1$ , а  $y \in \Pi_2$ , то в выписанной цепочке  $y$  заменяется на  $(xy)$  или соответственно на  $(yx)$ .

На основе этих операций цепочка должна порождаться следующим образом.

<sup>45</sup> Термин «валентность» (точнее: «синтаксическая валентность»), по-видимому, впервые был употреблен де Гроотом (см.: De Groot. Structurale syntaxis, стр. 160). У нас употребление этого термина, по-видимому, идет от С. Д. Кацнельсона.

Мы берем произвольное слово и совершаем в зависимости от того, подчиняет ли его пучок какой-нибудь другой пучок, операцию *A* или операцию *B* и переходим к связанному с ним слову. При этом каждый раз, когда мы совершаем операцию *A*, т. е. берем подчиненный пучок, мы запоминаем место, где в первый раз ее провели, и если данный пучок не является абсолютно-подчиняющим, то мы, проведя *n* раз операцию *A*, должны вернуться к данному месту и провести операцию *B*. Если же пучок является абсолютно-подчиняющим; то мы можем или сразу остановиться, или еще несколько раз провести операцию *A*, а затем остановиться.

Таким образом, этот процесс близок к тому, в котором мы, идя по дереву, строим систему составляющих (здесь система составляющих указывается скобками).

Рассмотрим процесс порождения некоторых предложений, порождавшихся ранее другими методами. Возьмем фразу *я вижу маленькую девочку*. Предложенная система порождения может начинаться с произвольного слова. Поэтому мы рассмотрим два разных способа.

Сначала выберем слово *я*. По извлечению из таблицы видно, что пучок (1-е л., ед. ч., И) — подчиненный, поэтому находим пучок (1-е л., ед. ч., наст. вр., перех.), подчиняющий этот пучок. Заменяем слово *я* на (*я вижу*). Берем в качестве обрабатываемого слово *вижу*. Так как соответствующий пучок абсолютно подчиняющий, то мы можем закончить процесс порождения на этом слове, а можем еще несколько раз применить правило *A*. Среди подчиненных слева имеется пучок (В, ЖО, ед. ч.), которому соответствует слово *девочку*. Заменяем (*я вижу*) на (*я (вижу девочку)*). Берем слово *девочку*, соответствующий пучок подчиняет справа пучок, соответствующий *маленькую*, заменяем (*я (вижу девочку)*) на (*я (вижу (маленькую девочку))*). Поскольку мы уже встретили абсолютно-подчиняющий пучок, то можем остановиться в любом месте, в частности, мы останавливаемся здесь.

Рассмотрим теперь способ порождения, начинающийся со слова *маленькую*. Поскольку соответствующий пучок — подчиненный, то мы можем применять лишь операцию *B*.

Заменяем *маленькую* на (*маленькую девочку*). Берем слово *девочку* и заменяем (*маленькую девочку*) на (*вижу (маленькую девочку)*) и т. д.

Мы видим, что данный способ порождения довольно близок к бесконтекстной грамматике (по-видимому, оба способа вообще эквивалентны), что в них одинаковую роль играют категории и что модель в некоторых случаях дает и членение по составляющим. Так, группа, получившаяся развертыванием по правилу *A* или *B*, трактуется как единая составляющая.

Хотя данная модель недостаточно исследована, но с лингвистической точки зрения она обладает некоторыми несомненными

преимуществами. Во-первых, то, что процесс порождения может начинаться с любого слова, по-видимому, приближает данную модель порождения к модели говорящего, ибо говорящий не только начинает фразу практически с любого слова (во всяком случае, это верно для подавляющего большинства слов в славянских языках), но и может произносить фразу именно из-за данного слова.

Во-вторых, в данной модели в отличие от модели непосредственно составляющих порядок слов не очень существен. Если не различать «подчинения справа» и «подчинения слева», а заменить в произвольном порядке одно слово сочетанием (*xy*) или (*yx*) (в правилах *A* и *B*), то мы получим не произвольные порядки слов, а как раз те, которые наиболее характерны для русского языка, например:

*я вижу девочку маленькую*  
*я маленькую девочку вижу*  
*я девочку маленькую вижу*  
*вижу девочку маленькую я*  
*вижу маленькую девочку я*

но не получим фраз:

*маленькую вижу девочку я*  
*маленькую я девочку вижу* и т. п.

Это особенно интересно с той точки зрения, что показывает связь данной модели с принципом проективности: она не порождает непроективных фраз (даже если необязательно фиксировать положение справа или слева относительно подчиняющего слова). Это обстоятельство не случайно: оно связано с тем, что в данной модели получается четкая система составляющих, отражаемая скобочной записью, а в этих условиях принцип проективности всегда действует.

Отметим, что принцип проективности связан с указанной и близкими ей в идейном отношении моделями и при аналитическом подходе.

Предположим, что мы решаем не задачу синтеза всех возможных фраз, а задачу, в некотором смысле обратную ей, а именно, мы знаем механизм порождения всего множества фраз и хотим восстановить последовательность правил, приводящую к порождению именно данного предложения. Здесь таблицы, в которых представлены все подчиняющие и подчиненные категориальные пучки, могут оказаться весьма ценными.

Построение грамматики валентностей, разработанное в ленинградской группе, занимающейся машинным переводом<sup>46</sup>, чрез-

<sup>46</sup> См.: Б. М. Лейкина. Некоторые аспекты характеристики валентностей. «Доклады на конференции по обработке информации, машинному

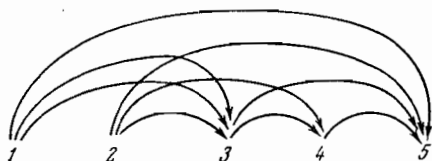
вычайно важно с практической точки зрения, но в ней вскрываются и некоторые существенные теоретические проблемы.

В частности, оказывается, что, если мы с помощью таблиц подобного вида будем устанавливать синтаксические связи в некоторой фразе, то получится ряд лишних (с точки зрения истинного анализа) связей. Так, если в таблице есть строка, указывающая на то, что любое существительное управляет существительным в родительном падеже, стоящем справа от него, и такое же управление имеет предлог *мимо*, то, например, в фразе

Проплывает желтый профиль  
мимо негра цвета кофе

мы установим зависимости:

*профиль*<sub>1</sub> *мимо*<sub>2</sub> *негра*<sub>3</sub> *цвета*<sub>4</sub> *кофе*<sub>5</sub>



И здесь принцип проективности помогает избавиться от большинства неверно установленных связей. Если считать, что все зависимости, нарушающие проективность, например

*профиль D негра*  
*профиль D цвета*  
*профиль D кофе*  
*негра D кофе*

неправомерно установлены, то остается гораздо меньше неверных связей:

*проплывает*<sub>1</sub> *профиль*<sub>2</sub> *мимо*<sub>3</sub> *негра*<sub>4</sub> *цвета*<sub>5</sub> *кофе*<sub>6</sub>



Остается лишь цепочка родительных падежей, но здесь от неправильных связей можно избавиться введением позиционных условий.

переводу и автоматическому чтению текста». М., Изд. ВИНТИ, 1961; С. Я. Ф и т а л о в. О моделировании синтаксиса в структурной лингвистике. «Проблемы структурной лингвистики». М, 1962, стр. 107—110.

• Оказалось, что и введение подобных позиционных условий не всегда обеспечивает правильные связи.

Н. А. Пашенко<sup>47</sup> приводит следующий показательный чешский пример:

studovat vliv stejných kosmických jevů na různé jedince

(‘изучить воздействие одинаковых космических условий на различных индивидумов’).

Для разбора этого предложения мы обязаны дать правило, по которому существительное управляет стоящей влево от него группой с предлогом на. Но тогда среди анализов будет и такой, в котором группа на jedince зависит от стоящего непосредственно перед ним существительного jevů.

Здесь можно выделить класс существительные вроде vliv ‘влияние’, обязательно требующих предлога на, а после того, как соответствующая валентность «насыщена», — не устанавливать других связей.

#### § 58. СОГЛАСОВАНИЕ, УПРАВЛЕНИЕ, ПРИМЫКАНИЕ

Введенное в § 18 отношение непосредственного подчинения есть объединение традиционных отношений согласования, управления и примыкания.

Естественно попытаться дать более точное определение этих понятий, поскольку эти понятия, как правило, довольно диффузны. Одна из попыток в этом направлении была предпринята О. С. Кулагиной; мы, однако, предложим здесь несколько иную систему определений, исходящую из общего понятия синтаксического подчинения (как мы видели, это понятие выводится из отношения «смысл фразы  $f$  включается в смысл фразы  $g$ »).

0. 7. 42. Говорим, что отношение  $D$  во фразе  $f$  и з б и р а т е л ь н о, если найдется такое  $y' \in \Gamma(y)$ , что ни в одной фразе  $f'$  не имеет места  $xDy'$ .

Примеры избирательных отношений:

*девочка D маленький* (имеется такое слово *маленький*, что ни в одной фразе нет отношения *девочка D маленький*).

*вижу D девочку* (невозможно ни в одной фразе *вижу девочку*).

Отношение названо избирательным, потому что управляющее слово выбирает определенные слова из парадигмы, причем в ней остаются такие слова, которые не могут быть подчинены данному слову.

<sup>47</sup> Н. А. Пашенко. Некоторые вопросы автоматического синтаксического анализа чешского научно-технического текста. «Научно-техническая информация», 1963, № 9, стр. 40.



Неизбирательными являются отношения *маленький D* очень (в парадигме *очень* вообще нет другого слова)

*вижу D* хорошо (в парадигме *хорошо* есть слово *лучше*, но возможно и *вижу D* *лучше*).

Этот пример показывает, что вполне можно определить примыкание следующим способом.

0. 7. 43. Говорим, что вхождение *y* примыкает к вхождению *x* во фразе *f*, если отношение  $xDu$  в *f* есть отношение неизбирательное (мы будем обозначать примыкание через  $D_3$ ).

Теперь разграничим понятия управления и согласования.

0. 7. 44. Говорим, что вхождение *x* управляет вхождением *y* во фразе *f* (мы будем обозначать управление через  $D_4$ ), если выполнены следующие условия:

а) имеет место  $xDu$  в *f*;

б) отношение *D* избирательно и

в) для любого  $x' \in \Gamma(x)$  найдется такая фраза *f'*, что  $x'Du$  в *f'*.

Последнее требование подчеркивает то обстоятельство, что свойство быть управляющим характеризует всю парадигму, например:

<i>вижу</i>	}	$D_4$ девочку
<i>видишь</i>		
<i>видя</i>		
<i>видящего</i>		
<i>дом</i>	}	$D_4$ отца
<i>дома</i>		
...		
<i>домами</i>		
<i>домов</i>		

Заметим, что наше определение вовсе не требует единственности управляемого слова и поэтому, например, в случаях типа *не вижу девочку* и *не вижу девочки* (ср. также примеры, приводимые в § 40) мы имеем дело опять-таки с управлением (отношение остается избирательным и выполняется третье требование). Между прочим, взяв все парадигмы, в которых найдется хотя бы одно слово, управляемое данным словом *x*, можно выбрать  $\Gamma(y)$ , содержащее максимальное число слов, управляемых словом *x*, и считать это число показателем слабости управления<sup>48</sup> (при этом в пределе слабое управление переходило бы в примыкание, т. е. неизбирательное подчинение).

<sup>48</sup> Ср. близкие идеи в кн.: «Основы построения описательной грамматики современного русского литературного языка». М., 1966, стр. 139—140.

Как показал Ю. Д. Апресян, разграничение сильного и слабого управления может быть лучше всего достигнуто, исходя из статистических принципов (см.: Ю. Д. Апресян. О сильном и слабом управлении. — ВЯ, 1964, № 3, стр. 32—49).

Согласование отличается от управления тем, что для него свойство  $v$ ) не выполняется, и поэтому вполне достаточно следующее определение.

**0. 7. 45.** Говорим, что вхождение  $y$  согласуется со вхождением  $x$  во фразе  $f$ , если имеет место  $xDu$  и это отношение не является ни управлением, ни примыканием.

Поскольку, однако, в славянских языках мы имеем дело с двумя разными видами согласования, а именно, со случаями типа *маленькая девочка*, где каждому слову из парадигмы *девочка* соответствует другое слово из парадигмы *маленькая*, и со случаями типа *гуляют D девочки*, где это свойство не выполняется.

Поэтому введем еще такие определения:

**0. 7. 46.** Отношение  $D$  называется сильным согласованием, если оно есть согласование и для любого  $x' \in \Gamma(x)$  и  $x' \neq x$  найдутся  $y' \in \Gamma(y)$  и  $y' \neq y$  и такая фраза  $f'$ , что  $x'Dy'$  (сильное согласование мы будем обозначать через  $D_1$ , а несильное через  $D_2$ ).

Таким образом, в случаях типа *девочка D маленькая* имеет место сильное согласование, а в случаях типа *гуляют D девочки* — слабое согласование.

**Примечание.** Определение 7.46 содержит известную идеализацию, а именно, оно предполагает, что омонимичные формы, например *маленькой* в *маленькой девочке* и *маленькой девочкой*, разграничены и трактуются как разные — можно вместо  $x' \neq x$  и  $y' \neq y$  потребовать, чтобы  $x'$  и  $x$  (и соответственно  $y'$  и  $y$ ) входили в разные категориальные пучки. Заметим, что требование  $x' \neq x$  (и соответственно  $y' \neq y$ ) существенно, ибо иначе сильное согласование оказалось бы частным случаем определенного выше управления, а нам хотелось разграничить эти понятия.

Как мы видим, наши определения довольно хорошо отражают действительное положение. Однако имеется один случай, где они сильно расходятся с интуицией. Речь идет о подчинении предлогов. По нашим определениям, во фразе *несогласен с директором* имеет место с  $D_4$  *директором*, т. е. управление предлогом существительного, что хорошо согласуется с интуицией, однако, с другой стороны, имеем место *согласен  $D_3$  с*, т. е. предлог с примыкает к глаголу. Между тем лингвистическая интуиция видит здесь управление (причем так называемое «сильное управление»). Здесь, однако, есть только один выход: мы уже указывали в ряде мест, что в синтаксическом анализе отношение подчинения целесообразно устанавливать не между словами, а между синтаксическими группами. Если считать, что *с директором* представляет собой единое целое (и при этом считать всю группу членом парадигмы слова *директор*), то тогда в соответствии с интуицией можно рассматривать этот случай как управление глагола предложной группой *с директором*.

Интересно рассмотреть с той же точки зрения отношение в сочетаниях типа *я вижу*. Вне зависимости от того, будем ли мы

считать, что слово *мы* входит в другую парадигму, чем *я*, как это было принято в § 26, или же в одну и ту же парадигму, все равно отношению *вижу D* я является согласованием, поскольку оно, во-первых, избирательно, а во-вторых, не является управлением. Это, правда, не противоречит традиционной точке зрения, но гораздо естественнее считать, что *я вижу* — вообще единая синтаксическая группа, внутри которой не устанавливается никакого отношения подчинения.

Мы уже указывали, что вряд ли целесообразно принимать разграничение типов подчинения по Пешковскому. Заметим, подход Пешковского не дал бы нам никакого типологического различия между языками, поскольку во всех языках есть отношение подчинения существительного или подчинения прилагательного другим словам. Иначе обстоит дело с введенными нами отношениями. А именно, в языках аморфных есть только примыкание; в таких агглютинативных языках, как тюркские, есть лишь примыкание и управление, а согласование ограничивается лишь отношением  $D_2$  (характерна при этом и необязательность такого согласования при связи существительного в именительном падеже множественного числа с формой 3-го лица глагола). В языках переходного типа между агглютинативным и флективным появляется и согласование типа  $D_1$  (причем сфера его уже, чем в флективных). Наиболее развито и разнообразно согласование именно в таких флективных языках, как славянские; причем это обстоятельство служит важным типологическим признаком флективных языков<sup>49</sup>.

Сформулированное нами разграничение согласования и управления может иметь ряд интересных приложений. Одно из них относится к уточнению вопроса об объеме глагольной парадигмы в славянских языках. В § 35 мы установили, что видовые формы, например *видишь* и *увидишь*, целесообразно считать относящимися к разным парадигмам, т. е. относить образование видовых форм не к области формообразования, а к области словообразования. Вопрос о залоговых формах мы в § 35 не касались. Объясняется это тем, что методами § 35 трудно установить, нужно ли относить *убивал* и *убит* к одной или к разным парадигмам. В пользу разделения парадигмы *убивал* на две разных парадигмы:  $\Gamma$  (*убивал*) и  $\Gamma$  (*убит*) говорит то обстоятельство, что только при такой трактовке в отношениях типа *убил D мальчика* можно усмотреть управление в строгом смысле. Одновременно управлением будет и отношение *убит D мальчиком*. В противном случае и та и другая связь подпадет под тип отношений  $D_2$ , т. е. будет формально такой же, как и в случаях вроде *убивает D мальчик*.

Итак, чисто синтаксический критерий приводит нас к необходимости признать, что образование залоговых форм относится

<sup>49</sup> Ср. также: V. Skalička. *Typ češtiny*. Praha, 1951, стр. 44—45.

не к ф ормообразованию, а к словообразованию <sup>50</sup>. В этом случае мы получаем возможность вполне строгого разграничения понятий управления и согласования.

Такое разграничение важно в связи с поставленной в начале параграфа задачей дифференциации отношений подчинения на схеме синтаксических отношений. У нас до сих пор во фразах типа *мальчик ругает девочку* предикативная связь не отличалась от комплетивной, что весьма существенно для анализа предложения, ибо в противном случае фразы *мальчик ругает девочку* и *девочка ругает мальчика* получают совершенно одинаковую трактовку. Теперь мы можем разграничить соответствующие типы связей.

Имеется, однако, весьма существенный недостаток в проведенном разграничении согласования (двух его видов), управления и примыкания. Связан он с тем, что в определении отражаются лишь факты парадигматики (выбор из парадигмы подчиненного слова), в то время как примыкание и слабое управление, с одной стороны, и сильное управление, с другой стороны, различаются весьма существенно и на синтагматической оси тем, что в первом случае существенна позиция подчиненного слова во фразе, в то время как во втором — связь между словами инвариантна относительно транспозиции, ср. *весьма восхищенный мальчик утром увидел цветущую ниву*, *восхищенный утром мальчик увидел весьма цветущую ниву* и *весьма восхищенный ниву увидел цветущую мальчик утром*, где первая и третья фразы находятся между собой в отношении транспозиции, хотя согласованные и сильно управляемые слова перемещены, в то время как вторая фраза дает совсем иные отношения из-за перемещения примыкающих и слабо управляемых слов. Связать предложенный нами парадигматический аспект с синтагматическим можно при помощи следующего определения.

0. 7. 47. Говорим, что во фразе  $f$  ограничена транспозиция, если в ней: а) для любого слова  $x$ , имеющего непустую валентность, во фразе  $f$  найдется слово  $y'$  такое, что  $y' \in \Gamma(y)$  и  $\Pi(y)$  входит в валентность слова  $x$ , и б) из двух слов  $y'$  и  $y''$  таких, что  $\Pi(y')$  и  $\Pi(y'')$  входят в валентность слова  $x$ , во фразе имеет место  $xDy'$ , если  $y'$  ближе к  $x$ , чем  $y''$  и  $xDy''$  в противном случае.

Посмотрим, как в таких условиях будут вести себя определенные нами типы связи.

---

<sup>50</sup> В этой связи уместно привести характеристику категории залога в русском языке как лежащей «уже на самой пограничной черте между морфологией, синтаксисом, лексикологией и фразеологией» (В. В. Виноградов. Современный русский язык, вып. II. М., 1938, стр. 469). Различие между активом и пассивом в русском языке, по мнению целого ряда авторов, относится не к морфологии, а к синтаксису.

При примыкании как неизбирательной связи каждое слово, имеющее непустую валентность, потенциально управляет примыкающим словом, если оно попадает в близкую позицию. При слабом управлении (ср. прим. на стр. 240) вероятность, что данная форма подчиняется и другому слову (если это позволено позицией), будет достаточно велика и, наконец, при сильном управлении и согласовании вероятность, что данное слово  $x$  подчиняет именно данную форму  $y'$  из парадигмы  $\Gamma(y)$ , будет минимальной. Таким образом при условиях, близких к действительному положению в славянских языках, оправдывается и разная зависимость данных типов связи от позиции слова.

Синтаксическое подчинение — не единственное отношение подчинения, которое можно установить между словами.

Так, И. А. Мельчук<sup>51</sup> предложил рассматривать наряду с синтаксическим подчинением также морфологическое и семантическое подчинение.

Поскольку в нашей работе мы не касались морфемного анализа, то понятие морфологического подчинения мы будем брать как интуитивное (тем более, что несмотря на трудность его строгого определения все лингвисты, по-видимому, однозначно решают вопросы морфологического подчинения, ибо содержательный смысл понятия вполне ясен: выбор формы слова  $x$  предопределен выбором формы слова  $y$ ). Что касается семантического подчинения, то это понятие трудно брать как интуитивное, и поэтому мы предложили свое определение, которое, может быть, и не целиком соответствует идее Мельчука, но судя по примерам, приводимым им, в большинстве случаев дает близкие результаты.

Идея нашего определения будет состоять в том, чтобы отождествить семантическое подчинение в некоторой фразе или ее части с синтаксическим подчинением в соответствующей ядерной фразе. Это приводит к следующей формулировке.

0.7.48. Говорим, что слово  $x$  семантически подчиняет слово  $y$  во фразе  $f$ , если а) в  $f$  имеет место  $xDy$  или  $yDx$  и б) существует такая последовательность фраз ядра  $g_1, \dots, g_n$ , что фраза  $g = g_1 \dots g_n$  находится в отношении трансформируемости с фразой  $f$  и одна из фраз ядра  $g_i$  содержит слова  $x'$  и  $y'$  такие, что  $x' \in \Gamma(x)$ ,  $y' \in \Gamma(y)$  и в  $g_i$  имеет место  $x'Dy'$  (как и во всех определениях, связанных с трансформируемостью, парадигма берется в самом широком смысле).

Интересно следующее. Хотя выбранные нами критерии и независимы (за исключением фраз ядра, на которых синтаксиче-

---

<sup>51</sup> И. А. Мельчук. Типы связи между элементами текста и типология языков. Материалы конференции «Актуальные вопросы современного языкознания и лингвистическое наследие Е. Д. Поливанова», т. 1. Самарканд, 1965, стр. 57—59.

ская зависимость совпадает с семантической), в целом как при определениях, принятых Мельчуком, так и при наших определениях можно установить ту типологическую особенность славянских языков, что в них в большинстве случаев все три типа зависимости совпадают, например:

*вижу мальчика*

*вижу хорошо*

или же семантическая зависимость совпадает с синтаксической при обратной морфологической зависимости, например:

*лист падал*

(причем и это несовпадение связано с той условностью, что личная форма глагола синтаксически подчиняет существительное в именительном падеже). Интересно, что в русском языке такие случаи, по-видимому, ограничиваются предикативным отношением, но в ряде других языков характеризуют и объектное отношение; таково по-видимому, объектное построение в венгерском языке:

szeretem Moszkvat

'люблю-ее' 'Москву',

где наличие объектной формы требует определенного имени, а наличие определенного имени требует объектной формы.

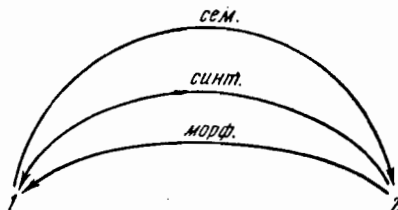
Во всех случаях этого типа имеет место согласование в слабом смысле, в то время как в первом случае имело место управление или примыкание.

В конструкциях, которые можно считать трансформированными, например в определительных, которые можно рассматривать как депрецицированные предикативные построения<sup>52</sup>, семантическое отношение не совпадает с морфологическим и синтаксическим, но направление двух последних отношений в определительных группах славянских языков, как правило, совпадает, например:

а) *маленький*<sub>1</sub> *мальчик*<sub>2</sub>

б) *сверкающий*<sub>1</sub> *луч*<sub>2</sub>

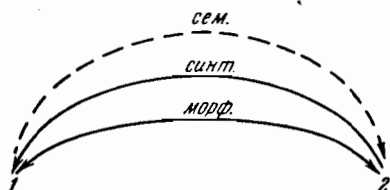
в) *книга*<sub>1</sub> *студента*<sub>2</sub>



<sup>52</sup> Ср. попытку формализации этого понятия: И. И. Ревзин. Формальный и семантический анализ синтаксических связей в языке. «Применение логики в науке и технике». М., 1960, стр. 133—135.

• Единственное и достаточно спорное исключение<sup>53</sup> составляют в славянских языках сочетания числительных *два*, *две* с существительными, например:

- а) *две*<sub>1</sub> *стены*<sub>2</sub>  
 б) *два*<sub>1</sub> *мальчика*<sub>2</sub>



Здесь категория рода слова *мальчика* требует формы *два*, а принадлежность слова *два* к особому подклассу числительных требует специальной формы существительного (оставляем в стороне вопрос о характеристике этой формы), т. е. имеет место взаимное морфологическое подчинение.

Типологически близкий и гораздо менее спорный случай представляет изафетное построение с родительным падежом в тюркских языках, например в казахском, ср.: *студент-тың кетаб-ы* 'студента' 'книга-его'. Здесь имеет место взаимное морфологическое подчинение при несовпадении направления семантического и синтаксического подчинения, но в отличие от приведенного русского примера с числительными речь идет о вполне регулярном явлении. И здесь интересно отметить, что в наших терминах эти случаи дают согласование в слабом смысле.

Совершенно нехарактерна для славянских языков обратная направленность морфологической и синтаксической зависимости, какая имеет, например, место в изафетных конструкциях персидского языка: *کتاب برادر* [кетаб-е бāрадār] 'книга-его брат — книга брата'.

Здесь синтаксически *bāradār* зависит от *кетаб*, а морфологически наоборот появление показателя *-e-* вызвано наличием слова *bāradār*.

Представляется, что дальнейшее уточнение этой схемы много даст для типологического анализа, но уже сейчас она довольно хорошо отражает синтаксическую специфику славянских языков.

В заключение этого параграфа отметим одно отличие наших определений от схемы Мельчука. В последней включаются в рассмотрение и отношения между членами предложных групп, например в группе *позади дома* Мельчук констатирует совпадение направления всех трех отношений. В наших определениях мы

<sup>53</sup> Спорность состоит в том, что неясно, можно ли считать относящимися к ядру фразы типа *стен было две*.

лишены возможности установить семантическое отношение между членами сочетания *позади дома*. Разумеется, эти трудности формального характера можно было бы обойти, видоизменив систему определений. Но поскольку, как мы пытались показать в ряде других мест, синтаксические отношения целесообразно устанавливать не между отдельными словами, а между целыми группами, то мы считали возможным и здесь не рассматривать отдельно предложные группы.



## О ВОЗМОЖНОСТЯХ МОДЕЛИРОВАНИЯ СЕМАНТИКИ

§ 59. НЕКОТОРЫЕ ТРУДНОСТИ ПОСТРОЕНИЯ  
СЕМАНТИЧЕСКИХ МОДЕЛЕЙ

Семантические факты и в первую очередь явления категоризации, которым придавалось столь существенное значение в § 1, почти не нашли отражения в книге (за исключением явлений грамматической категоризации, т. е. распределения слов по грамматическим категориям, ср. в особенности разделы, посвященные роду и падежу существительных). Такое ограничение неправомерно сузило аппарат, применяемый для типологии в том смысле, как она была определена в § 5. Известно, что именно на семантическом уровне можно обнаружить наиболее существенные сходства и несходства (даже когда речь идет о круге близкородственных языков).

Мы не исследовали, однако, фактов семантики не потому, что не могут быть предложены семантические модели, достаточно интересные по своим идеям. Однако до сих пор эти модели строятся почти на голом месте. Если теория грамматических моделей может опереться на хорошо разработанный применительно к ряду конкретных языков материал, причем выяснен основной круг категорий и их противопоставлений, то в области семантики соответствующий фактический материал не собран и не расклассифицирован, а в руководствах и монографиях по лексикологии и семасиологии в лучшем случае можно найти лишь разрозненные примеры, уровень систематизации которых далеко отстает от самых начальных руководств по грамматике.

Таким образом, дело не только в отсутствии фактического материала, но и в том, что имеющийся (хотя и незначительный) материал не подвергался серьезному теоретическому осмыслению. Если в области грамматики моделирование (как мы подчеркивали в § 3) уже давно имело место, хотя оно и было неявным и не всегда последовательным и задача состоит в уточнении старых и построении на их основе новых, более совершенных моделей, то в области семантики такая работа никогда не была проведена. Вопреки утверждениям противников структурализма традиционное языкознание во всяком случае в лице своих достаточно авторитетных представителей никогда серьезно не занималось семантикой (и

поэтому упреки в игнорировании смысловой стороны языка явно направляются не по адресу).

Поэтому основная задача состоит не столько в конструировании семантических моделей, сколько в последовательном описании конкретной смысловой структуры соответствующих языков. Отсутствие фактического материала приводит и к тому, что весьма трудно выяснить характер исходных понятий, на основе которых могут успешно строиться семантические модели.

Поэтому мы хотели бы ограничиться лишь указаниями на возможность введения некоторых семантических понятий, которые на абстрактном семиотическом уровне соответствуют грамматическим, ибо если удастся построить такую систему понятий, то весь формальный аппарат можно было бы заимствовать из уже разработанных моделей.

Однако и эта скромная задача связана с огромными трудностями, в связи с чем все последующие замечания следует рассматривать как носящие сугубо предварительный характер (поэтому, в частности, эта глава не будет содержать в отличие от других ни определений, ни теорем).

#### § 60. О ПОНЯТИИ СЕМАНТИЧЕСКИ ПРАВИЛЬНОЙ ФРАЗЫ

Первый вопрос, на котором мы хотели бы остановиться, это различие между понятиями «реальной фразы», «семантически правильной фразы», «логически осмысленной фразы» и грамматически правильной фразы. Грамматически правильную фразу нельзя отождествить с семантически правильной: это очевидно из примеров фраз, приводимых в § 17. Можно ли отождествить семантически правильную фразу с «реальной фразой», как это сделано Небеским в модели, обсуждавшейся в § 17? Даже если исходить из предпосылки, неявно принятой нами, хотя и специально неговоренной (ибо в соответствующем месте нас этот вопрос просто не интересовал), а именно что любая «реальная фраза» семантически правильна (т. е. мы исходили из того, что все тексты, дающие исходный материал для исследования, осмыслены), все равно следует признать, что такая идентификация вряд ли обоснована.

Мы постараемся сейчас показать, что обратное включение неоправдано, т. е., что не всякая семантически правильная фраза является фактически фразой, встретившейся в тексте обозримой длины.

Во-первых, следует отметить, что никакой реальный текст на естественном языке не устроен так, что он исчерпывающе описывает некоторый мир в смысле Лейбница-Витгенштейна<sup>1</sup> (т. е.

<sup>1</sup> Ср. понятие «описания состояния» у Карнапа как дающего «полное описание возможного состояния универсума индивидов относительно всех

некоторую совокупность вещей и отношений между ними или ситуаций). Скорее наоборот любой текст на естественном языке избегает исчерпывающего описания мира. Единственное исключение, известное автору, это тексты упражнений в учебниках иностранных языков. Однако именно эти тексты поражают своей неестественностью и представляют поэтому благодарный материал для пародирования. Показателен следующий рассказ Ионеску, описывающего, как он пришел к идее создания своей пьесы «Лысая певица» (*La cantatrice chauve*) в результате чтения учебника английского языка:

«... Я старательно переписывал фразы из моего учебника, чтобы выучить их наизусть и вот, перечитывая их, я познакомился не с английским языком, а с удивительными истинами: что в неделе семь дней (это я знал и раньше) или что пол находится снизу, а потолок наверху (может быть, я это тоже знал, но забыл, во всяком случае я об этом никогда не думал, а здесь мне эта мысль показалась весьма странной, хотя и несомненно верной)... После универсальных истин, автор учебника раскрыл мне частные истины, причем для этого автор, вдохновленный несомненно методом Платона, выражал их посредством диалога. С 3-го урока появились два персонажа, не могу сказать: реальных или вымышленных: мистер и миссис Смит, двое англичан. К моему восхищению миссис Смит сообщила своему супругу, что у них несколько детей, что они живут в окрестностях Лондона, что их фамилия Смит, что мистер Смит — служащий, что у них есть прислуга Мэри, также англичанка, что вот уже 20 лет они имеют друзей по фамилии Мартин... Я позволю себе привлечь Ваше внимание к тому факту, что утверждения миссис Смит носили абсолютно бесспорный аксиоматический характер и что автор моего учебника целиком следовал методу Декарта, ибо наиболее удивительным было то, что в поисках истины он действовал весьма методично... В 5-м уроке... за элементарными истинами последовали более сложные: на утверждение «в деревне спокойнее чем в большом городе» следовал ответ «да, но в городе более плотное население и, кроме того, там больше магазинов»<sup>2</sup>.

Итак, полный набор семантически правильных фраз резко расходится с тем, что принято считать текстом. Язык избегает полноты описания мира. Это достигается путем соотношения речи с контекстом (внеязыковым) высказывания, склеивания предложений, неполных перечислений, эллипсисов всякого рода и т. п.<sup>3</sup>

свойств и отношений, выраженных предикатами этой системы» (Р. Карнап а п. Значение и необходимость. М., 1959, стр. 38). Там же Карнап добавляет, что «описания состояния представляют возможные миры Лейбница или возможные состояния вещей по Витгенштейну» (стр. 39).

<sup>2</sup> E. Ionesco. *La tragédie de langage*. «Notes et contre-notes». Paris, 1962, стр. 155.

<sup>3</sup> Эта особенность естественного языка была отмечена уже Я. Линдбахом, которого во многих отношениях можно считать предшественником со-

Если реальные фразы суть фразы, действительно встретившиеся в тексте, то следует констатировать, что не всякая семантически правильная фраза есть реальная фраза.

Во-вторых, семантически правильные фразы могут не встретиться ни в одном реальном тексте потому, что выражаемая ими истина еще не пришла в голову ни одному носителю языка (таковы весьма важные для развития науки предложения, в которых без всякого нарушения семантической правильности формируются новые факты и гипотезы).

Семантически правильные фразы нельзя отождествить и с (логически) осмысленными фразами. Дело в том, что в этом случае семантически правильными следовало бы признать все фразы, которые являются либо истинными, либо ложными. Но даже если ограничиться утвердительными предложениями, имеются семантически правильные фразы такие, что соответствующие им предложения ни истинны, ни ложны. Таковы, например, фразы типа *человек существует, кентавр не существует* или же неполные (с логической точки зрения) высказывания (например, *дождь полезен*).

По-видимому, семантическая правильность связана с ощущением того, что в данной ситуации соответствующее высказывание уместно, или точнее мотивировано.

Понятие мотивировки было выдвинуто русскими формалистами в связи с анализом литературного произведения<sup>4</sup>. Представляется однако, что это понятие имеет общесемиотическое значение уже на уровне анализа отдельной фразы. Семантически правильной фраза может стать под действием контекста (соответствующий тип мотивировки назывался русскими формалистами композиционной мотивировкой). Так, неоднократно приводившаяся фраза *идея яростно спит*, являющаяся семантически неправильной, может быть мотивирована в следующем контексте<sup>5</sup>: *Идея яростно спит / Ворочается во сне / Идея в висках стучит / Нашептывая мне*. Этот тип мотивировки весьма распространен в поэзии. В сущности, однако, и упомянутые выше фразы, формулирующие новые истины в науке, мотивированы с семантической точки зрения именно композиционно.

---

временной семиотики (см.: Я. Линдбах. Принципы философского языка. Гг., 1916, стр. 39, 89 и сл.). Линдбах заметил, что примеры относительно полного языка встречаются в былинах, ср.

Садка день не зовут на почестен пир,  
Другой не зовут на почестен пир,  
И третий не зовут на почестен пир.

<sup>4</sup> См.: Б. В. Томашевский. Теория литературы. М., 1928.

<sup>5</sup> Ср.: И. И. Ревзин. Некоторые трудности при построении семантических моделей для естественных языков. «Симпозиум по структурному изучению знаковых систем. Тезисы докладов». М., 1962.

В обычном языке мотивировка осуществляется другими средствами, а именно сочетаются между собой лишь такие слова, семантические категории которых допускают соответствующее сочетание. Эти категории, накладываясь на грамматические, образуют своего рода фильтр, с помощью которого здравый смысл<sup>6</sup> отвергает предложения типа *идея яростно спит* (если они не мотивированы композиционно).

Поэтому для естественного языка вопрос о семантически правильных фразах также связан с вопросом о семантических категориях как вопрос о грамматически правильных цепочках связан с вопросом о грамматических категориях (ср. § 24).

Вопрос о порождении семантически правильных фраз на основе реальных фраз в сущности уже поставлен в работах Хомского<sup>7</sup> и развит в работе Катца и Фодора<sup>8</sup>. Порождение можно было бы осуществить так. Нужно разбить грамматические категории на подкатегории, например категорию существительных на подкатегории абстрактных (абстр.) и конкретных (конкр.), категорию глаголов на подкатегории физического состояния (физ.) и абстрактного состояния (абстр.), категорию наречий на категорию активного признака (акт.) и пассивного признака (пасс.) и т. п., тогда грамматика:

Предложение → (Ж, ед. ч., И, конкр.) + (3-е л., ед. ч. наст., физ.)

Предложение → (Ж, ед. ч., И, абстр.) + (3-е л. ед. ч., наст., абстр.)

(3-е л., ед. ч., наст., физ.) → (наречие, пасс.) + (3-е л., ед. ч., наст., физ.)

(Ж, ед. ч., И., конкр.) → *девочка*

(Ж, ед. ч., И, абстр.) → *идея*

(3-е л., ед. ч., наст. вр., физ.) → *спит*

(3-е л., ед. ч., наст. вр., абстр.) → *существует*

(наречие, пасс.) → *спокойно*

совместима с языком, содержащим фразы *девочка спокойно спит* и *идея существует*, и при этом не производит семантически неправильных фраз вида *идея яростно спит*.

Хомский предлагает, правда, считать соответствующие подкатегории грамматическими и соответственно порождаемые фразы грамматически правильными (на более высоком категориальном

<sup>6</sup> В концепции Б. В. Томашевского такому способу соответствует реалистическая мотивировка.

<sup>7</sup> Н. Хомский. Несколько методологических замечаний о порождающей грамматике. Перев. с англ. — ВЯ, 1962, № 4, стр. 119 и сл.

<sup>8</sup> J. Katz, J. Fodor. The Structure of a Semantic Theory. «Language», 39, 1963, стр. 170—210.

уровне, связывая с этим идею уровней грамматической правильности). Однако нам представляется, что удобнее рассматривать соответствующие категории как семантические, ибо это более соответствует лингвистической традиции, различающей семантический и грамматический уровни абстракции.

Возникает, однако, вопрос, каким образом получить необходимый набор семантических категорий. Подход Хомского предполагает эмпирическое сопоставление порождаемых грамматикой фраз и выделение новой подкатегории всякий раз, когда одновременно порождаются и какие-либо семантически неправильные фразы. Это, разумеется, весьма надежный, но с другой стороны, трудоемкий путь, который хотелось бы как-нибудь сократить.

Такого сокращения можно было бы добиться, если бы у нас была бы заранее определенная гипотеза о возможном репертуаре семантических категорий. Ниже мы выскажем такую гипотезу, но прежде всего рассмотрим вопрос о том, как изучается семантика грамматических категорий.

#### § 61. СЕМАНТИКА ГРАММАТИЧЕСКИХ КАТЕГОРИЙ

В главе VI мы изучали формальную структуру грамматических категорий. Нас в сущности не интересовал вопрос, что такое падеж или род или число и т. п., мы определяли лишь, что значит высказывание: два слова  $x$  и  $y$  относятся к одному числу, роду или падежу. Очевидно, что на этом, вообще говоря, не только обычное лингвистическое описание, но и моделирование остановиться не может. Но хотя вся лингвистическая практика всегда состояла в определении значений той или иной формы и соответственно категории, только в последнее время вопрос о значении категорий был поставлен в форме, допускающей дальнейшую формализацию.

Мы имеем в виду ряд замечательных работ Р. Якобсона по русскому языку и в первую очередь его работу о падежах<sup>9</sup>. В этих работах, сколько может судить автор, впервые была сформулирована задача определения семантики категории как совокупности значений небольшого числа дифференциальных семантических признаков. Так для русских падежей (мы ограничимся здесь шестипадежной системой) Р. Якобсон выделил три весьма общих семантических признака: направленности, периферийности и объемности, значения которых следующим образом распределяются по падежам:

<sup>9</sup> R. Jakobson. Beitrag zur allgemeinen Kasuslehre. — TCLP, 6, 1936, стр. 240—288.

Падеж	Признак		
	направлен- ность	периферий- ность	объемность
И	—	—	—
Р	—	—	+
Д	+	+	—
В	+	—	—
Тв.	—	+	—
Пр.	—	+	+

Весьма общий и абстрактный характер этих признаков часто вызывал недоумение исследователей, пытавшихся применить их к конкретным фактам (или даже осмыслить объяснение отдельных фактов, даваемых Якобсоном)<sup>10</sup>. Однако основная идея (может быть, и не до конца осознанная самим автором) состоит не в объяснении отдельных употреблений падежа, а в различении значения форм двух падежей в тождественном окружении, а эту задачу система Якобсона, по-видимому, выполняет (а если понадобится, то всегда можно ввести и некоторые дополнительные признаки).

Для нас здесь особенно существенно, что каждый из признаков Якобсона, взятый в отдельности, вовсе не обязательно ограничивается сферой имени существительного. Как подчеркивает сам автор, признак направленности характеризует не только падежные противопоставления, но и залоговые противопоставления глагола<sup>11</sup>, а признак объемности, связанный с понятием предела, характеризует и видовую корреляцию глагола в русском языке<sup>12</sup> (ср. одинаковое воздействие отрицания на этот признак в формах имени и глагола: *возьми стул — не бери стула*).

Таким образом, репертуар семантических дифференциальных признаков может быть единым для разных категориальных функций, а каждая функция (определяющая систему категорий рода, падежа и т. п.) характеризуется лишь определенной конфигурацией признаков (здесь слово конфигурация употреблено, разумеется, не в терминологическом смысле).

Что касается категорий рода, то здесь тем же Якобсоном были предложены также бинарные признаки (типа женскость —

<sup>10</sup> Ср. обсуждение доклада Р. Якобсона (Р. Якобсон. Морфологические наблюдения над русским склонением. s'-Gravenhage, 1958). — В кн: «IV Международный съезд славистов. Материалы дискуссии», т. II. М., 1962, стр. 39—48.

<sup>11</sup> Там же, стр. 254.

<sup>12</sup> Там же, стр. 255.

неженскость); если сюда добавить противопоставление: наличие биологического пола — отсутствие его, то мы легко придем к обычной тернарной классификации. Как мы видели, для славянских языков необходимы еще противопоставления личности — величность и одушевленность — неодушевленность. Интересно, однако, что несколько расширив эти признаки, можно получить набор их, дающий возможность описать семантические противопоставления, которые имеются во всех языках, грамматически выражающих род.<sup>13</sup>

Мы видим, что понятие категоризации и утверждение о разной категоризации действительности, выдвинутые в § 1, нуждаются в одном существенном уточнении: разным, по-видимому, является не набор семантических дифференциальных признаков категорий, а те или иные способы их объединения в категории. В связи с этим становится понятной и возможность взаимопонимания людей, говорящих на разных языках<sup>14</sup>.

Но если некоторые объединения признаков не реализуются в данном языке грамматически, то, по-видимому, такие объединения могут считаться семантическими категориями. В связи с этим соображением мы и выдвинем в следующем параграфе определенный принцип формирования семантических категорий, который, как представляется, может сократить чисто эмпирический путь их выявления.

## § 62. ОБ ОДНОМ ВОЗМОЖНОМ СПОСОБЕ ФОРМИРОВАНИЯ СЕМАНТИЧЕСКИХ КАТЕГОРИЙ

Как мы уже говорили, семантические категории, приписываемые словам, до сих пор вводятся *ad hoc*, причем не существует формальной процедуры, проверяющей правильность выделения соответствующих категорий.

Поэтому мы предлагаем следующий эмпирический принцип.<sup>15</sup> Мы будем считать признак  $\pi$  семантически категориальным для любого естественного языка, если существует хотя бы один есте-

---

<sup>13</sup> По материалам, которые см.: L. A d a m. *Du genre dans diverses langues*. Paris, 1883; R o u e n. *Die nominalen Klassifikationssysteme in den Sprachen der Erde*. Wien, 1929.

<sup>14</sup> Высказанные здесь и далее в § 63 соображения, как представляется, еще сильнее ограничивают значимость известной гипотезы Сэпира—Уорфа, чем те, которые высказаны в § 15 кн. И. И. Ревзина и В. Ю. Розенцвейга «Основы общего и машинного перевода» (М., 1964, стр. 68—75). Факты различной категоризации действительности, на которых основана эта гипотеза, получают при этом более простое объяснение.

<sup>15</sup> См.: О. Г. Карпинская. К методике обнаружения некоторых универсалий плана содержания. «Конференция по проблемам изучения универсальных и ареальных свойств языков. Тезисы докладов». М., 1966, стр. 33—34.



ственный язык, в котором данный признак, входит в совокупность, определяющую некоторую грамматическую категорию.

Этот принцип возвращает нас к так называемым «понятийным категориям» О. Есперсена<sup>16</sup> и отчасти И. И. Мещанинова<sup>17</sup>, незаслуженно забытым в последнее время. Применение его может дать полезные результаты, если считать, что

а) грамматические категории выделяются объективно, а сведение каждой к небольшому набору признаков выполнимо;

б) набор семантических категорий тождественен для всех языков мира и связан с особенностями человеческой психики.

Первое предположение достаточно хорошо обосновано всей грамматической традицией, а ко второму мы еще вернемся.

Можно заметить, что даже оставаясь в пределах одного языка, можно извлечь из анализа категорий, представленных в одних разрядах слов, семантические признаки, полезные для анализа семантических категорий других слов того же языка.

Так в одной из самых важных последних работ, посвященных анализу семантики<sup>18</sup>, этот принцип неявно присутствует при выделении целого ряда семантических операций и семантических признаков. Например, выделяется семантическая операция конверсии (Conv), т. е. называния того же отношения, но взятого в другом направлении: Conv (*выше*) = *ниже*, Conv (*включать*) = *входить в*, Conv (*любить*) = *нравиться*, Conv (*вперед*) = *позади* и т. п.

При этом делается замечание, что «наряду с приводимыми здесь лексическими конверсиями имеется регулярный способ образования конверсивов — пассив (*резать — резаться, нагревать — нагреваться* и т. п.)»<sup>19</sup>.

Это верно и для других вводимых авторами операций, например, Saus 'сделать так, чтобы', хотя ими и не оговаривается (ср. регулярный способ образования каузативных глаголов) или Magn *очень* 'высокая степень' (ср. образование превосходной степени, и в особенности значение элатива, которое имеют большинство слов русского языка на *-ейший, -айший*).

По-видимому, весьма существенны и косвенные свидетельства данных грамматики. Так, в § 47 мы видели, что группа «оценочных» прилагательных (*высокий — низкий, хороший — плохой* и т. п.) образует в русском языке особый абстрактный род. Существенной семантической особенностью этой группы является, с одной

<sup>16</sup> О. Есперсен. Философия грамматики. Перев. с англ., М., 1958, стр. 57—60.

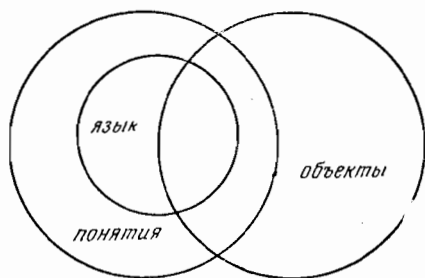
<sup>17</sup> И. И. Мещанинов. Части речи и члены предложения. М., 1945.

<sup>18</sup> А. К. Жолковский, И. А. Мельчук. О возможном методе и инструментах семантического анализа. «Научно-техническая информация», № 6, 1965, стр. 23—28.

<sup>19</sup> Там же, стр. 25.

стороны, то, что они группируются в антонимичные пары, а с другой стороны, то, что они все образуют координатную сетку для ориентации человека. Пополнив эту группу другими прилагательными, обладающими теми же чертами, мы получим группу, чрезвычайно удобную для семантической классификации и чрезвычайно важную для общей теории категоризации<sup>20</sup>.

Это согласуется и с гипотезой, выдвинутой специалистом по проблемам детской речи Е. Леннебергом<sup>21</sup>. На основании рассмотрения большого числа экспериментальных данных Леннеберг приходит к выводу, что соединение слова и понятия есть результат развития, которому предшествует врожденная (т. е. передаваемая генетически) система референциальных понятий типа *близкий — далекий*, *выше — ниже* и т. п., построенных по бинарному принципу. По Леннебергу соотношение между миром объектов и миром понятий передается схемой



Причем на этой схеме образование устойчивых связей между миром объектов и миром понятий отражается именно тем, что язык обязательно затрагивает сферу пересечения их.

Конечно эта гипотеза нуждается в проверке, однако показательно, что к тем же выводам приходят авторы, исследующие вопрос о месте семантических категорий в порождающей грамматике (ср. упомянутый доклад М. Бирвиша).

Таким образом, можно предположить, что существует универсальный набор семантических признаков, единый для всех языков мира и связанный с особенностями устройства психики человека (так же как существует единый набор фонетических признаков, связанный с биологическими особенностями человека), а типология должна иметь дело с различными комбинациями этих семантических признаков. Методика такого анализа

<sup>20</sup> Это обстоятельство было блестяще раскрыто М. Бирвишем в докладе на симпозиуме по семиотике (август 1965, Варшава).

<sup>21</sup> E. N. L e n n e b e r g. The relationship of language to the formation of concepts. «Synthese», 1962, 14, № 2—3, стр. 103—109.

для одной частной -подсистемы была разработана применительно к славянским языкам Н. И. Толстым<sup>22</sup>.

Если считать, что каждому слову или точнее синтаксической группе, образуемой полнозначным словом и примыкающими к нему служебными словами соответствует определенный пучок семантических признаков, то вполне могут быть использованы для типологии некоторые предложенные ранее модели.

Так, аналогично понятию фонологически полной подсистемы<sup>23</sup> может быть введено понятие семантически полной системы. Пусть каждому слову  $x$  поставлен в соответствие набор семантических признаков  $\pi_1, \pi_2 \dots, \pi_n$ . Назовем два слова соседними, если они различаются только тем, что в соответствующем наборе одно из значений признака  $\pi_i$  заменено на другое. Система называется полной, если любые два слова можно соединить цепочкой, в которой каждые два слова соседние.

Интересно, что мы в несколько других терминах дали то определение, которое предложил Вейнрейх для семантически непрерывной системы<sup>24</sup>.

Можно было бы перенести в семантику и ряд других моделей<sup>25</sup>, но, как мы уже говорили, дело не в моделях, а в фактическом материале, а для славянских языков, кроме упомянутой работы Н. И. Толстого, мы не нашли ни одной работы, из которой можно было бы извлечь более или менее систематизированные данные.

Из-за недостатка фактического материала в области лексикологии, мы возьмем случай, промежуточный между грамматикой и лексикологией, а именно суффиксальное словообразование<sup>26</sup>, и попробуем применить к нему общие принципы, обычно выдвигаемые в связи с построением фонологических моделей.

Применительно к продуктивному словообразовательному суффиксам задача может быть сформулирована следующим образом: каждому суффиксу (для простоты можно считать омонимичные суффиксы уже разграниченными) необходимо приписать такую совокупность семантических признаков, в которой а) каждый признак не должен быть связанным по отношению к соответствующей

<sup>22</sup> Н. И. Толстой. Из опытов типологического исследования славянского словарного состава. — ВЯ, 1963, № 1, стр. 29—45.

<sup>23</sup> См.: И. И. Ревзин. Модели языка, стр. 33.

<sup>24</sup> См.: U. Weinreich. Lexicographic definition in descriptive semantics. «Problems in lexicography, International Journal of American Linguistics», Vol. 28, № 2, part. IV, стр. 34—35.

<sup>25</sup> Общая идея перенесения понятий фонологических и грамматических моделей в семантику принадлежит Вяч. Вс. Иванову, см.: Вяч. Вс. Иванов. Понятие нейтрализации в морфологии и лексике. «Бюллетень объединения по проблемам машинного перевода», № 5, 1957, стр. 56.

<sup>26</sup> Мы будем пользоваться при этом идеями и фактами, изложенными в неопубликованной работе О. Г. Карпинской «К типологии словообразовательных систем в славянских языках».

щей основе (например не повторять семантических категорий, приписываемых основе и б) совокупность семантических признаков должна быть выбрана так, чтобы два несинонимических слова, образованные от одной и той же основы с разными суффиксами получили разную запись на уровне признаков (ср. § 13).

Первое условие как правило нарушается в традиционных словообразовательных классификациях: Так считается, что имена с суффиксом *-iĭ* обозначают а) происхождение из той или иной местности (*костромич*, *москвич*), б) отчество (*Ильич*, *Лукич*, *Кузьмич*)<sup>27</sup>, хотя очевидно, что эти семантические признаки целиком связаны с основой.

Считают также, что суффикс *-n'ak* «обозначает молодую рошу, группу деревьев кустов»<sup>28</sup>, любопытно, что при этом оговаривается, что это имеет место «в сочетании с именными основами, обозначающими деревья» (ср. *сосняк*, *ивняк* и т. д.). Таких примеров можно привести сколько угодно. С другой стороны, ряд тонких замечаний в цитированной работе позволяет сформулировать более общие признаки, чем мы в дальнейшем и воспользуемся.

Что касается второго принципа, то мы предложим следующие семантические признаки, которые в отличие от иногда выдвигаемых логических признаков<sup>29</sup> мы будем называть **к в а л и ф и к а т и в н ы м и**.

а) Интенсивность — неинтенсивность.

Особенность суффикса *-aĭ* в некоторых славянских языках в том, что он не только указывает на наличие у обозначаемого лица свойства, указанного основой, но подчеркивает (в отличие от других суффиксов деятеля) **и з б ы т о ч н о е, и н т е н с и в н о е** присутствие этого свойства. Например, в русском языке: *силач* — человек необычайной силы мышц; *ловкач* — ловкий парень, расторопный, годный на все, ловкий человек; *носач* — долгоносый (о человеке, животном, посуде, орудии) *брюхач* — брюхастый, толстопузый.

В каждом из этих слов суффикс *-aĭ* выполняет три функции: 1) переводит производящую основу в разряд имен деятеля, 2) указывает на обладание производным именем признака, указанного в основе, 3) подчеркивает, что этот признак присутствует в большем количестве по сравнению с обычным. Именно эта последняя черта суффикса *-aĭ* является в нем основной по семантике, проти-

<sup>27</sup> См.: В. В. Виноградов. Современный русский язык. М., 1938, стр. 49.

<sup>28</sup> Там же, стр. 60.

<sup>29</sup> В теории славянского словообразования идея понятийных (логических) признаков последовательно развивалась Дорошевским (см.: W. Doroszewski. Zarys klasyfikacji znaczeniowej sufiksów w językach słowiańskich. «Sbornik Praci I Sjezdu Slovanskych Filologu». Praha, 1929; см. также: Он же [выступление на дискуссии] «IV Международный съезд славистов, материалы дискуссии», П. М., 1962, стр. 128).

Этот автор выделяет три признака: 1) субъект—предикат, 2) пассив—актив и 3) потенциальность—актуальность.

вопоставляя его другим суффиксам деятеля и иногда выступая ведущей опорой в образовании имени. Так, *носач*, *брюхач* возможны только потому, что суффикс *-ащ* указывает в них на наличие у владельца выдающегося носа и выдающегося брюха (иначе такое имя никого бы не выделяло, потому что у всех людей или во всяком случае у большинства есть носы и животы).

Суффикс *-ащ* редко образует новые слова в русском языке. Но и в них находит проявление признак интенсивности, окрашивающий суффикс, хотя само название может уже выглядеть метафорически: *крылач* — спец. вентилятор (т. е. с большими крыльями); *трюкач* — тот, кто хорошо умеет делать трюки, человек, любящий выделять ловкие проделки, *затем*, неожиданные поступки, производящие большое впечатление<sup>30</sup>.

На интенсивности суффикса *-ащ* строит Хлебников свое знаменитое: *О, расмейтесь, смехачи*. Здесь интересно, что не только подбор лексических корней создает интенсивную атмосферу смеха, но и сам выбор суффикса способствует сгущению этой атмосферы (*смехачи*, *смеячи*).

Аналогическое значение интенсивности выделялось другими исследователями в русском языке и для суффикса *-ап*, который обозначает лицо «по какому-нибудь преобладающему признаку, по какому-нибудь чрезмерно развитому свойству»<sup>31</sup> (В. В. Виноградов выделяет в этом суффиксе и другое значение — экспрессивности, о котором ниже).

Интересно, что для суффикса *-ака* еще Миклошич отметил указание на энергичное проявление или усиление действия, по которому обозначается лицо<sup>32</sup>.

Показательно, однако, что такое же значение интенсивности отмечалось применительно к другим категориям, например, именам действия, ср. как описывает суффикс- *jož* в русском языке В. В. Виноградов: «... носящий отпечаток разговорной энергичной экспрессии и вместе с тем подчеркивающий интенсивность самого действия: *делёж*, *грабёж*, *кутёж*, *платёж*»<sup>33</sup>.

Интересно, что суффикс *-еж* описывается в болгарском языке как нейтральный<sup>34</sup>, но зато усилительность приписывается суффиксу *-ба* (*-ит-ба*) с помощью которого «образуются некоторые существительные, обозначающие действие с оттенком, продолжительности или усиления: *борба*, *веселба*, *делба*, *жалба*, *служба*, *стрелба*, *сватба*, *нареда*, *молба*, *просба*, *кражба*, *беритба*, *вършитба*, *женитба*, *коситба*, *ловитба* и др.»<sup>35</sup>

<sup>30</sup> С. И. Ожегов. Словарь современного русского языка. М., 1955.

<sup>31</sup> В. В. Виноградов. Современный русский язык. Грамматическое учение о слове, II, стр. 56.

<sup>32</sup> См. там же, стр. 29.

<sup>33</sup> См. там же, стр. 52.

<sup>34</sup> Л. Андрейчин. Грамматика болгарского языка. Перев. с болгарского. М., 1949, стр. 72.

<sup>35</sup> Там же.

Заметим, что признак интенсивности сам по себе не ограничен именами: как грамматический он проявляет себя в формах элитива прилагательных (*крупнейший* = *очень, весьма крупный*), а на лексическом уровне выражается весьма различно<sup>36</sup>.

Примечания: 1. Введенное нами бинарное противопоставление, несомненно, является идеализацией. Можно было бы привести некоторые соображения и в пользу тернарного противопоставления: а) высокая, б) нейтральная и в) пониженная интенсивности. Так, суффикс *-š'ik* в русском языке указывает по сравнению с *-tel'* на интенсивное обладание свойством или интенсивное занятие чем-то, ср.: *покупщик* — *покупатель*, *носильщик* — *носитель*, *наборщик* — *набиратель*, *зажимщик* — *зажиматель* и т. п. Наоборот, в паре *-tel'* и *-jes*, *-es* по интенсивности маркирован суффикс *-tel'* (ср. *владелец* — *владелец*). Таким образом, суффикс *-es* в сравнении с суффиксом *-š'ik* указывает на ослабленную интенсивность. Однако нам существенно именно противопоставление ненейтральности по интенсивности и нейтральности.

2. Представляется чрезвычайно знаменательным тот факт, что интенсивными могут быть как правило лишь суффиксы мужского рода, который вообще с точки зрения словообразования резко противопоставлен женскому и среднему (ср. обратное противопоставление в системе склонения, где женский род резко противопоставлен неженским). Можно думать, что семантическое содержание родов, стертое в грамматике, остается живым в словообразовании, отражаясь на принципах организации его структуры.

#### б) Предельность — непредельность.

Если сравнить значение суффикса *-ač* в русском языке с соответствующим материально суффиксом *-ač* в южнославянских языках, то легко установить, что в них этот суффикс выражает значение временной ограниченности действия, которое сближает имена деятеля на *-ač* с действительными причастиями настоящего времени.

Например, в сербском языке: *вешач* — 1) вешающий что-л., 2) палач, вешающий; *давач* — дающий, давший что-либо человек; *питач* — человек, спрашивающий что-либо<sup>37</sup>.

В болгарском языке: *мияч* — человек, который моет, умывает; *записвач* — тот, кто записывает; *запитвач* — 1) (разг.) тот, кто спрашивает, задает вопросы, 2) лицо, которое вносит запрос<sup>38</sup>.

В македонском языке: *берач* — собирающий плоды; *делач* — делающий что-либо; *пяч* — поющий<sup>39</sup>.

<sup>36</sup> Этот признак играет важную роль в той формализации семантических фактов, которая предложена И. А. Мельчуком и А. К. Жолковским («О возможном методе и инструментах семантического анализа», стр. 25).

<sup>37</sup> Сербско-хорватско-русский словарь под ред. И. И. Толстого. М., 1957.

<sup>38</sup> Болгарско-русский словарь под ред. С. Б. Бернштейна. М., 1947; Сава Чукалов. Пълен русско-болгарски речник. София, 1958. — О возможности обозначить именем на *-ač* лицо по непостоянному признаку см.: Е. А. Захаревич. Производные основы со значением лица в современном болгарском литературном языке. Сб. «Вопросы грамматики болгарского литературного языка». М., 1959.

<sup>39</sup> В. Магков. *Nomina actionis* w języku macedonskim. «Poradnik językowy», 1965, zes. 6, стр. 241. — Б. Марков показывает, как при закреп-

В каждой из этих систем, суффикс *-ač* указывает на непостоянство, временную ограниченность признака, положенного в основу названия. Характерно, что на русский язык такие имена переводятся именно причастиями. Об идентичной роли имен на *-ač* и причастий говорит также и их взаимная компенсация: в сербском, болгарском языках, где имена на *-ač* с маркированностью по признаку временной ограниченности чрезвычайно распространены, причастия употребляются гораздо реже и наоборот — в русском, где отсутствует суффикс имени деятеля с маркированностью по признаку временной ограниченности, роль причастий более значительна.

В русском языке у имен деятеля подобное противопоставление отсутствует, но зато противопоставление по временной ограниченности/неограниченности весьма характерно для имен действия, ср. *выигрши* — *выигривание*, *щелчок* — *щелкание* и т. п.

Было бы, однако, неверно останавливаться на данной ступени абстракции, ибо значение времени (т. е. ограниченности во времени) привносится глагольным характером основы (идея времени связана с глаголом). Если же выделить оставшееся значение, то им будет идея предела, ограничения. Этот признак, по-видимому, близок признаку объемности у Якобсона, а также близок понятию актуализации (той роли, которая приписывается артиклю в западноевропейских языках). Применительно к именным основам оно реализуется как противопоставление большей конкретности (ограничение объема в пространстве) или наоборот абстрактности.

В русском языке по этому признаку противопоставлены субстантивированное прилагательное (*сейте разумное, доброе, вечное*) и имя свойства (*разумность, доброта, вечность*).

В болгарском же языке, где, как мы видели, нейтрализуется разница между причастием и именем деятеля, аналогичное положение имеет место и для имен свойства<sup>40</sup>, ср. отношения имеющие место в болгарском языке между суффиксами *-ota* и *-ina*, *-ost*, *-stv* и *-šcin*. По этому поводу Л. Андрейчин пишет: «В большинстве случаев наряду с каждым существительным на *-ota* — существует и другое с суффиксами *-ина*, *-ост*, обозначающее то же качество, в более прямом и конкретном смысле, например: *леснота* — *леснина*, *доброта* — *добрина*, *лекота* — *лекост*, *мекота* — *мекост* . . .». И еще: «. . . суффикс *-щина* приближается по своему значению к суффиксу *-ство*, от которого отличается только большей конкретностью: *юнащина* — *юначество*, *чове-*

---

лении имени на *-ač*, т. е. снятии признака предельности. оно переходит в *non professionalia*.

<sup>40</sup> Может быть, это связано и с меньшим распространением субстантивированных прилагательных в южнославянских языках, о чем говорит И. Леков («Общность и многообразие в грамматический строй на славянские языки». София, 1958, стр. 76).

щина — човечност, дивайщине — дивачество . . . и др.»<sup>41</sup>. Элементы такого противопоставления видны и в русском языке с лексикализацией членов этого противопоставления (ср. *широта* — *ширина*).

Интересно, что признак предельности/непредельности может участвовать в глагольной системе наряду с признаком интенсивности/неинтенсивности в качестве целиком грамматикализованного. Именно таково, по-видимому, положение в языке ки-конго<sup>42</sup> (один из языков банту), где имеются а) специальная форма, указывающая на интенсивность действия (диалект Майомбе: *sumba* 'покупать', *sumbidila* 'покупать все разом, «сильно» покупать'), б) специальная форма, указывающая на непредельность действия (*sumbanga* 'постоянно покупать') и в) форма, совмещающая эти значения (*sumbidilanga* 'много раз, все время покупать много').

Таким образом, сами признаки не специфичны для суффиксов, специфична лишь их определенная конфигурация.

**П р и м е ч а н и е.** Как и в случае первого противопоставления, здесь можно было бы ввести тернарную классификацию: а) ограниченность значительным отрезком (например, при проявлении этого признака как временная ограниченность указание на действие, происходящее в течение длительного периода, ср. значение суффикса *-on'* в *тихоня*: человек, о котором подчеркивается, что он всегда тихий) б) нейтральность по этому признаку и в) ограниченность малым отрезком (указание на действие происходящее в данный момент, например, указанное значение *-ac* в болгарском языке). Но мы и здесь ограничиваемся бинарным противопоставлением: маркированности по признаку указания на предел или его отсутствие и нейтральности по этому признаку.

#### в) Субъективность — несубъективность.

Мы уже говорили, что в ряде суффиксов интенсивность связана с другим признаком, который иногда характеризуют как экспрессивность, а иногда как субъективную оценку. Очевидно, что признак субъективной оценки может выступать и без признака интенсивности (например, *квартирка*), и наоборот признак интенсивности без субъективной оценки (*дележ*, *грабеж*, *беготня* и т. п.).

Если первые два противопоставления нужно было характеризовать более полно, то третье противопоставление описано достаточно хорошо<sup>43</sup> (и даже выяснено отличие этого признака от обычных значений). Поэтому мы не будем приводить специально новых примеров.

**П р и м е ч а н и е.** И в этом признаке возможна тернарная классификация, а именно: а) субъективно увеличивающая оценка, б) нейтральная оценка и в) субъективно уменьшающая оценка, ср. обычное деление субъек-

<sup>41</sup> Л. Андрейчин. Указ. соч., стр. 73.

<sup>42</sup> L. de Clecq. Le verbe Kikongo. «Congo». Juin, 1935, t. II, № 1.

<sup>43</sup> См., например: С. С. П л я м о в а т а я. Размерно-оценочные имена существительные в современном русском языке. М., 1961.



ективных суффиксов на увеличительные и уменьшительные (заметим, что чисто увеличительными или уменьшительными бывают лишь имена вещей, орудий, мест, т. е. такие, в которых наши квалификативные признаки не проявляются). Однако для нас и здесь (как и в остальных противопоставлениях) важно наличие самой субъективности или отсутствие ее.

Если сравнивать славянские словообразовательные системы с точки зрения их материальных компонентов их следует признать очень близкими или даже тождественными друг другу<sup>44</sup>. Это находит свое естественное объяснение в генетической общности славянских языков и их происхождении из одного источника — общеславянского языка. Но структурные типологические различия между ними, безусловно, имеются, и это чувствует любую личность, знающей более одного славянского языка. Однако традиционная классификационная схема суффиксальных значений не позволяет выявить их, она негибка и оказывается то слишком широкой, то слишком узкой, так что реальные специфические различия славянских словообразовательных систем ускользают от выявления. С этой точки зрения анализ по квалификативным признакам представляется более перспективным. Выделяя общие для славянских языков классы имен деятеля, действия и состояния и т. д., мы можем поставить вопрос, каковы взаимоотношения между суффиксами внутри каждого из этих классов и какие отношения складываются между самими суффиксальными классами. И здесь специфический набор квалификативных признаков, возможность их сочетаемости друг с другом, а также маркированность специальными суффиксами отдельных наборов квалификативных признаков выступают теми типологическими характеристиками, по которым можно надеяться выявить различия славянских словообразовательных систем.

Нам хотелось бы привести лишь один пример. Для простоты мы выделим всего три класса суффиксов: а) имен деятеля, б) имен действия и свойства и в) имен вещей (включая сюда имена орудий и имена места). Суффиксы имен деятеля в русском языке участвуют в противопоставлениях по всем трем признакам, суффиксы имен действия и свойства участвуют в двух противопоставлениях (по предельности и интенсивности) и, наконец, имена вещей участвуют только в противопоставлении по субъективности, которое, к тому же, как правило, реализуется для этого класса в качестве противопоставления по размеру (увеличительность или уменьшительность). Таким образом, класс имен деятеля выступает содержательным центром русской словообразовательной системы, в то время как два других класса (имена действия

---

<sup>44</sup> Как показывает Ив. Леков («Словообразовательни склонности на славянските езици». София, 1958), различия здесь проходят по числу суффиксов, их продуктивности и т. д., но они не касаются структурных отношений систем.

и имена вещей) не имеют между собой общих признаков и беднее в признаковом отношении (особенно имена вещей), чем класс имен деятеля. Наоборот, на уровне форм выражения центром является класс имен вещей, ибо он вбирает в себя опустошенные суффиксы из двух других классов. Таким образом, здесь подтверждается та закономерность, что, чем сильнее маркирована языковая единица, чем богаче она по содержанию, тем специфичнее сфера ее употребления и наоборот.

То, что описанный характер участия отдельных классов в тех или иных квалификативных оппозициях специфичен для русского языка, подтверждается сравнением с болгарским языком, где картина существенно изменяется. В частности, имена деятеля участвуют всего в двух противопоставлениях (по предельности и субъективности), имена действия и свойства также в двух (по предельности и интенсивности), а имена вещей также, как и в русском, участвуют лишь в одном противопоставлении.

Недостаток места не позволяет показать, что указанная методика может быть распространена на отдельные суффиксы и зафиксированные наборы значений и что здесь также могут быть вскрыты существенные типологические различия. В качестве примера вновь возьмем суффикс *-ač*.

Язык	Признак		
	предельность	интенсивность	субъективность
Русский	—	+	—
Польский	—	—	—
Сербский, болгарский	+	—	—

Таким образом, польский суффикс соответствует по семантике русским *-tel'*, *-šćik*, а сербский и болгарский, скорее, соответствуют русскому причастию.

Представляется, что подобные сопоставления могли бы оказаться полезными в типологии. Но до сих пор, как уже было указано, для систематического их проведения и дальнейшей формализации нет достаточного фактического материала.

### § 63. ПАРАДИГМА КАК СЕМАНТИЧЕСКОЕ СЛОВО

Мы уже обсудили два исходных понятия семантической модели, а именно понятия семантически правильной фразы и семантической категории. Этого, однако, недостаточно для перенесения построенного аппарата в область семантики, поскольку нам необходимо еще по крайней мере понятие исходной единицы (аналог словоформы в грамматических моделях). Можно, разумеется, взять ту же словоформу, в качестве исходного понятия. Однако в этом случае, мы, по-видимому, не получим лингвистически значимых результатов.

Посмотрим, например, какой лингвистической интерпретации следует ожидать от понятия взаимозамещаемости на множестве семантически правильных фраз. Например, слова *безразличие*, *равнодушие* и *влечение*, по-видимому, будут взаимозамещаемыми на этом множестве, а любое из этих слов и слово *отдохновение* не будут взаимозамещаемыми (ср. *отдохновение от трудов*). Это говорит о том, что класс эквивалентности относительно множества семантически правильных фраз можно, по-видимому, интерпретировать как семантическое поле.

Однако для того, чтобы такая интерпретация была бы значимой, хотелось бы, чтобы в тот же класс, что *безразличие*, *равнодушие*, *вдохновение* входили бы слова *апатия*, *вялость*, *восторг*, а в тот же класс, что и *отдохновение*, входило бы слово *отдых*. Но на уровне фраз мы этого не получим из-за разного согласования.

Поэтому для семантических моделей следует брать не словоформы, а парадигмы и (взаимо)замещаемость рассматривать не относительно фраз, а относительно *G*-образов этих фраз.

Таким образом, вся парадигма слова участвует в семантической модели в качестве исходной единицы. Эту роль исходного элемента целесообразно приписать парадигме в узком смысле как совокупности форм словоизменения.

«Семантической парадигмой» следует считать совокупность форм словообразования, т. е. некоторое словообразовательное гнездо. Заметим, что именно такой парадигмой мы пользовались во всех тех местах, где обсуждалось отношение трансформируемости, связанное в какой-то мере с идеей семантической инвариантности. Поэтому все модели, используемые в связи с трансформируемостью можно с равным успехом считать и семантическими.

### § 64. УРОВНИ ЯЗЫКА И ТИПЫ МОДЕЛЕЙ

Укрупнение исходного элемента при переходе от грамматического уровня к семантическому (вместо словоформы единицей становится вся парадигма), укрупнение — происходящее в парадигматике языка — интересно сопоставить с укрупнением исходного элемента на синтагматической оси: при переходе от фоне-

тического уровня к морфологическому единицей становится словоформа, а при переходе от морфологического к синтаксическому исходным элементом становится синтаксическая группа.

Мы уже отмечали в § 56, что морфологический и синтаксический уровни естественно описывать разными моделями порождения: морфологический при помощи автоматной грамматики, а синтаксический — при помощи грамматики непосредственно составляющих. Поскольку, однако автоматная грамматика столь же хорошо описывает порождение словосочетаний, включающих не только служебные слова, но также и полнозначные зависимые члены (определения к существительному, обстояательства и дополнения к глаголу), то представляется возможным выдвинуть следующую гипотезу.

Можно представить себе дело так, что один и тот же текст синтезируется одним и тем же устройством альтернативно одним из двух способов: или синтезируются отдельные словоформы (или синтаксические группы), каждой из которых ставится в соответствие отдельный деизигнат, или же сразу же синтезируются целые словосочетания и лишь такое словосочетание в целом получает некоторый деизигнат. Описанная вторая возможность могла бы быть использована для моделирования автоматизма повседневной речи и некоторых других функциональных стилей («подъязыков»), когда отдельные элементы речи и синтагматические отношения между ними «не выплывают в светлое поле сознания». Еще в 20-е годы участники ОПОЯЗ'а и Московского лингвистического кружка показали принципиальное различие между автоматизмом повседневного употребления языка, связанным с бессознательными механизмами его осуществления, с одной стороны, и той особенностью поэтического языка, что он использует речь сознательно преобразованную, стремящуюся, наоборот, вывести адресата из автоматизма восприятия и дать ему почувствовать значение каждого элемента языка <sup>45</sup>.

Разумеется самодовлеющая поэтическая речь (или соответствующий пересмотр основ языка науки в определенные периоды), с одной стороны, и абсолютно автоматизованная речь «людоедки Элочки» у Ильфа и Петрова, с другой стороны, суть предельные ситуации — в остальной же речи реализуются, как правило, оба типа построения сообщения.

В этой связи гипотеза о возможности построения одного и того же сообщения из более мелких и более крупных блоков приобретает особый интерес. В главе V, изучая модели деятельности лингвиста при анализе текста как средство типологического иссле-

<sup>45</sup> Имеется в виду в первую очередь теория «остранения», выдвинутая В. Щкловским, см.: В. Шкловский. Искусство как прием. «Поэтика. Сборники по теории поэтического языка», I—II. Пг., 1919, стр. 101—104. — Ср. также теорию Р. Якобсона, по которой поэтическая речь характеризуется установкой не на сообщение, а на код.

дования языков, мы видели, что хотя различные языки совершенно по разному организуют соотношение вещественных значений («парадигмы») и формальных значений («семейства»), упорядоченность, симметричность соответствующих парадигм («формальная однородность») лишь повышается при укрупнении исходного отрезка. Отсюда следует, что порождение целых словосочетаний в известном смысле более экономно. Правда, при этом исходный словарь весьма существенно увеличивается (в принципе вместо множества слов мы берем множество пар или троек слов) и для всего языка в целом мог бы достичь фантастических размеров. Но все дело в том, что именно те подязыки («функциональные стили»), для которых характерен наибольший автоматизм (сфера бытового общения с одной стороны и газетная и деловая речь, с другой), отличаются сравнительно малым запасом слов и высокой степенью устойчивости сочетаний.

Отметим, что все грамматические категории, которые мы изучали в данной работе, а именно грамматический род и грамматический падеж, а также аналогично устроенные категории других частей речи (ср. главу VI), могут быть вскрыты лишь одним способом, а именно путем пробы на взаимозамещаемость отдельных слов внутри словосочетаний (во всех определениях основную роль играло понятие «семейства»). Это означает, что если брать каждое словосочетание как целое, то грамматические категории отдельных слов уже не могут быть вскрыты, хотя, как мы видели, они могут участвовать в порождении словосочетаний (ср. интерпретацию состояний в § 49). Такое положение в какой-то мере отражает то обстоятельство, что в автоматизированной речи категории никак не осознаются. С другой стороны предельно расчлененная поэтическая речь делает иногда упор и на грамматические категории слов. Еще более существенен тот механизм образования по аналогии в поэтической речи, о котором мы говорили в § 19 и который приводит к расширению возможностей сочетания слов, а тем самым обращает внимание на категории каждого отдельного слова. В пользу возможности функционирования словосочетаний как исходных единиц (в одной из двух возможных альтернатив порождения) говорит и тот факт, что в естественной речи любое словосочетание более или менее связано фразеологически (причем степень свободы определяется, в первую очередь статистически). Фразеологическая связанность в терминах теории моделей лучше всего может быть описана схемой с конечным числом состояний: после данного слова имеется строго определенное (причем небольшое) число возможностей употребления других слов. Наоборот в поэтической речи происходит, как известно, преодоление, а иногда и полное нарушение фразеологических связей. Модель, которая производит клише, не разлагая их на составные части, но снабжена параллельно действующим устройством, анализирующим (синтезирующим) в случае необходимости соответству-

ющие группы, разлагая их на составные части и приписывая каждой свое значение, а всей группе определенную иерархию этих значений, представляется актуальной для исследования проблем речевого общения и его эффективности. Представление каждой фразы, как имеющей четкую структуру, характерное для Хомского и его последователей, имеет, по-видимому, своим источником чрезмерную интеллектуализацию речевых фактов<sup>46</sup> (это отражается и в отборе материала, отражающего нормы школьной грамматики — ср. § 51) — рассмотрение каждого сочетания в речи как отражающего творческую активность мысли. На уровне отдельных слов очевидно, что говорящий (или соответственно слушающий) не синтезирует (анализирует) слово вроде *пастух* из двух элементов и что он может поступать двояким образом в случае слов типа *раститель*, трактуя их либо как целое, либо как возникающую синтагму. Мы предлагаем не только распространить подобную концепцию и на уровень словосочетаний, но и связать с каждым способом особую модель порождения.

Перейдем теперь к предложениям. Мы видели, что автоматная грамматика может с успехом производить и некоторые простые предложения. По указанной концепции это означает, что некоторые предложения тоже можно представить как готовые речения, не подвергаемые анализу<sup>47</sup>. В целом однако, как мы видели, простые предложения моделируются лучше всего грамматикой, дающей каждому предложению ясное членение, соответствующее в первую очередь его субъектно-предикатной структуре. С этим согласуется тот семантический факт, что простые предложения выполняют уже не номинативную функцию, а функцию коммуникативную, сообщая нечто о чем-то. При этом простые предложения, как правило, соответствуют атомарным высказываниям.

Наконец, как мы видели, сложные предложения лучше всего описываются трансформационной моделью, в которой исходными элементами являются простые предложения.

Изложенная здесь концепция, основанная на сознательном сближении порождающих грамматик, с одной стороны, и синтеза, с другой стороны, ставит в соответствие каждой из моделей порождения в иерархии, установленной Н. Хомским уже в работе «Синтаксические структуры» (автоматные грамматики, модели непосредственно составляющих, трансформационные грамматики), определенную лингвистическую реальность и тем самым, как представляется, позволяет яснее увидеть преимущества, связанные с идеей порождающих грамматик.

<sup>46</sup> По-видимому, неслучайны недавние заявления Хомского о его идейной близости с картезианской лингвистикой и рационалистической грамматикой Пор-Руаяля, ср.: N. Chomsky. Current Issues in Linguistic Theory. The Hague, 1964, стр. 8, 15, 16.

<sup>47</sup> Ср.: Н. Хоккетт. Грамматика для слушающего. Перев. с англ. «Новое в лингвистике», вып. IV. М., 1965, стр. 142.

Наконец, отметим, что в представленной иерархии грамматик центральное место занимают грамматики непосредственно составляющих. Об этом свидетельствует не только то обстоятельство, что порождение этим методом составляет существенную часть трансформационной грамматики (порождение «ядерных предложений»), но и то, что трансформационная грамматика обслуживает наиболее интеллектуализованную речь, в какой-то мере чужеродную для ядра языка. Существенно и то, что правила развертывания в грамматике составляющих носят наиболее общий характер, будучи по своей структуре одинаковыми для всех славянских (да и не только славянских) языков. Они столь же универсальны, как и те семантические признаки, из которых формируются грамматические категории отдельных языков. В то же время правила воплощения (т. е. по нашей концепции переход на уровень автоматной грамматики, а именно уровень оформления слов и словосочетаний) обнаруживают большие расхождения не только с точки зрения материальной, что вполне естественно, но и с точки зрения структуры. Это и проявилось в главах IV—VI, где были зафиксированы наибольшие различия между славянскими языками.

## ПРИЛОЖЕНИЕ

### ДОКАЗАТЕЛЬСТВА ИЛИ ССЫЛКИ НА ДОКАЗАТЕЛЬСТВА ТЕОРЕМ

**Т. 2. 1.** См. любое пособие по абстрактной алгебре или теории множеств.

**Т. 2. 2.** См.: В. А. Успенский. Лекции о вычислимых функциях. М., 1960, стр. 143—149.

**Т. 3. 1.** Исходя из реальных и запрещенных цепочек можно породить процедурой § 20 в числе фраз запрещенную цепочку и, кроме того, существуют цепочки, которые не запрещены и не являются фразами.

**Доказательство.** Пусть  $V = \{x, y, z\}$ , множество реальных фраз  $\alpha = \{xyz, zx, yx\}$ , множество запрещенных цепочек  $\beta = \{xzy\}$ . Тогда  $z$  дистрибутивно близко  $y$ . Цепочка  $xuz$  есть фраза: тогда фразами будут и цепочки  $xzx$  и  $xzy$ , но последняя цепочка запрещена. Цепочка  $zx$  не запрещена и не фраза. Теорема доказана.

**Т. 3. 2.** См. доказательство теоремы 5.2 в «Моделях языка», стр. 179—180.

**Т. 3. 3.** Если  $K_1 \Rightarrow K_2$ , то любой контекст, входящий в  $K_1$ , является более емким, чем любой контекст, входящий в  $K_2$ .

**Доказательство.** Пусть  $(f_1, g_1) \in K_1$  и  $(f_2, g_2) \in K_2$ . Надо доказать, что  $(f_1, g_1) \Rightarrow (f_2, g_2)$ . По определению существуют такие контексты  $(f_3, g_3) \in K_1$  и  $(f_4, g_4) \in K_2$ , что  $(f_3, g_3) \Rightarrow (f_4, g_4)$  (1). Раз  $(f_1, g_1)$  и  $(f_3, g_3)$  входят в один контекстный класс, то  $(f_1, g_1) \Rightarrow (f_3, g_3)$  (2). Аналогично  $(f_4, g_4) \Rightarrow (f_2, g_2)$  (3). По отношению большей емкости транзитивно и поэтому сопоставляя (2), (1) и (3) — в этом порядке — мы получаем, что  $(f_1, g_1) \Rightarrow (f_2, g_2)$ . Теорема доказана.

**Т. 3. 4.** а) Пусть  $\Delta$  — правильный фрагмент. Тогда  $\bar{\Delta}$  есть правильный фрагмент б) если  $\Delta_1$  и  $\Delta_2$  правильные фрагменты, то  $\Delta_1 \cup \Delta_2$  и  $\Delta_1 \cap \Delta_2$  суть правильные фрагменты.

**Доказательство.** Пусть  $x \sim y$  относительно  $\theta$ .

а) Пусть  $x \neq y$  в  $\bar{\Delta}$ . Это означает, что существуют такие цепочки  $f$  и  $g$ , что  $fxg \in \bar{\Delta}$ , а  $fyg \notin \bar{\Delta}$  или наоборот  $fyg \in \bar{\Delta}$ , а  $fxg \notin \bar{\Delta}$ . Предположим, что верно первое, тогда  $fyg \in \Delta$ , а  $fxg \notin \Delta$ , аналогично и в обратном случае. Отсюда следует, что  $x \neq y$  относительно  $\Delta$ , что противоречит условию, ибо из правильности фрагмента следует, что  $x \sim y$  относительно  $\Delta$ .

б) Пусть  $\Delta_1$  и  $\Delta_2$  суть правильные фрагменты. Предположим, что  $\Delta_1 \cup \Delta_2$  не есть правильный фрагмент. Тогда существуют такие цепочки  $f$  и  $g$ , что  $fxg \in \Delta_1 \cup \Delta_2$  (1), а  $fyg \notin \Delta_1 \cup \Delta_2$  (2) (или наоборот). Из (1) следует, что  $fxg \in \Delta_1$  (3) или  $fxg \in \Delta_2$  (4). Пусть верно (3). Но из (2) следует, что  $fyg \notin \Delta_1$  а, стало быть,  $x \neq y$  относительно  $\Delta_1$ . Но  $x \sim y$  относительно  $\theta$  и значит фрагмент  $\Delta_1$  — неправильный.

в) Аналогично можно доказать, что  $\Delta_1 \cap \Delta_2$  есть правильный фрагмент. В самом деле пусть  $fxg \in \Delta_1 \cap \Delta_2$  (1), а  $fyg \notin \Delta_1 \cap \Delta_2$  (2). Из (2) следует, что



$fyg \notin \Delta_1$  (3) или  $fyg \notin \Delta_2$  (4). Но  $fxg \in \Delta_1$  и  $fxg \in \Delta_2$ . Это противоречит условию, что  $x \sim y$  относительно  $\Delta_1$  и  $\Delta_2$ . Теорема доказана.

Т. 3. 5. а) Для любого  $k$  множество фраз длины, не превосходящей  $k$ , есть правильный фрагмент.

б) Для любых  $k$  семейств  $S_1, \dots, S_k$  множество фраз, в которых не содержится слов из  $S_1, \dots, S_k$ , есть правильный фрагмент.

Доказательство. а) Пусть  $x \sim y$  относительно  $\theta$ . Обозначим через  $\Delta^{(k)}$  множество правильных фраз длины, не превосходящей  $k$ . Пусть  $fxg \in \Delta^{(k)}$ , тогда  $fyg \in \theta$ , а поскольку длина не меняется при замене  $x$  на  $y$ , то  $fyg \in \Delta^{(k)}$ .

б) Нам достаточно доказать, что фрагмент, не содержащий слов из некоторого фиксированного семейства  $S_1$  есть правильный, ибо дальше мы можем просто взять пересечение  $k$  таких фрагментов. Итак, пусть  $\Delta_1$  — фрагмент, содержащий все фразы, кроме тех, которые содержат слова из  $S_1$ . Пусть  $x \sim y$  относительно  $\theta$  и  $fxg \in \Delta_1$ . Предположим, что  $fyg \notin \Delta_1$ . Тогда в этой фразе хотя бы одно слово  $z$  входит в  $S_1$ . Это слово не входит ни в  $f$ , ни в  $g$ , ибо иначе и  $fxg$  не принадлежало бы  $\Delta_1$ . Итак,  $z = y$ , но  $y \in S(x)$  и стало быть  $fxg \notin \Delta_1$ . Теорема доказана.

Т. 3. 6. см. доказательство леммы 2 в «Моделях языка», стр. 173.

Т. 5. 1. См. доказательство теоремы 4.2 в «Моделях языка», стр. 175—176.

Т. 5. 3. См. доказательство теоремы 12 в кн.: S. Marcus. *Linguistică matematică*, cap. VI, стр. 179.

Т. 5. 4. См. доказательство теоремы 4.3. в «Моделях языка», стр. 177.

Т. 5. 5. См. доказательство леммы 4 в «Моделях языка», стр. 185—186.

Т. 5. 6. См. доказательство теоремы 6 в работе: И. И. Ревзин. Распространение теоретико-множественной модели на язык с омонимией. «Научно-техническая информация», 1965, № 5, стр. 36.

Т. 5. 7. См. там же доказательство теоремы 7, стр. 36.

Т. 5. 8. См. там же доказательство теоремы 8, стр. 36.

Т. 5. 9. Следует из рассуждений, проводимых в указанной статье.

Т. 5. 10. См. доказательство теоремы 19 в кн.: S. Marcus. *Linguistică matematică*, стр. 183—184.

Т. 5. 11. Аморфный язык полностью однороден тогда и только тогда, когда для любых  $x$  и  $y$  из  $x \Rightarrow y$  следует  $x \in S(y)$ .

Доказательство: Пусть у нас имеется некоторый аморфный язык. Предположим, что он полностью однороден и что имеет место  $x \Rightarrow y$ . По определению полностью однородного языка в таком языке имеет место следующее; существуют  $x' \in \Gamma(x)$  и  $y' \in \Gamma(y)$  такие, что  $x' \in S(y')$ . Но поскольку язык аморфный  $x' = x$  и  $y' = y$ , а стало быть  $x \in S(y)$ . Итак, необходимость условия доказана. Пусть теперь язык аморфен и для любых  $x$  и  $y$  из  $x \Rightarrow y$  следует  $x \in S(y)$ . Всякий аморфный язык формально однороден. Покажем, что он и полностью однороден. В самом деле пусть  $x \Rightarrow y$ . Но тогда по условию  $x \in S(y)$ . Теорема доказана.

Т. 5. 12. См. доказательство теоремы в работе: И. И. Ревзин. Некоторые формальные особенности парадигмы глагола. «Исследования по структурной типологии». М., 1963, стр. 102 (следует лишь обратить внимание на разницу терминологии: в указанной работе предполагается, что каждая фраза является фразой с полной трансформацией).

Т. 5. 13. По существу это утверждение равносильно теореме 5.12 в «Моделях языка» (см. приведенное там на стр. 185—186 доказательство).

Т. 6. 1. Пусть категория  $x_1$  является одновременно парадигматической и синтагматической и пусть  $x \in x_1$ . Тогда  $Ra(x) \subseteq x_1$ .

Доказательство. Пусть  $y \in Ra(x)$ . Если длина  $\Gamma$ - $S$ -цепи равна двум, то  $y \in \Gamma(x)$  или  $y \in S(x)$ . В обоих случаях  $y \in x_1$ . Пусть для  $\Gamma$ - $S$ -цепи, состоящей из  $n-1$  членов верно, что  $y_{n-1} \in x_1$ , но  $y_{n-1} \in \Gamma(y)$  или  $y_{n-1} \in S(y)$ , а значит  $y \in x_1$ . Теорема доказана.

**Т. 6. 2.** Если язык формально однороден, что в нем все синтагматические категории равномерно распределены по парадигмам, а все парадигматические категории равномерно распределены по семействам.

**Доказательство.** Пусть язык однороден,  $x \in Ra(y)$ ,  $z \in \Gamma(x) \cap \chi_1$ , где  $\chi_1$  — синтагматическая категория. По теореме 5. 3. существует слово  $u$  такое, что  $u \in \Gamma(y) \cap S(x)$ . Но тогда  $x \in S(u) \cap \Gamma(z)$  и в силу однородности существует  $w$  такое, что  $w \in S(z) \cap \Gamma(u)$ . Раз  $w \in S(z)$ , а  $z \in \chi_1$ , то  $w$  входит в силу синтагматичности категории в  $\chi_1$ . При этом  $w \in \Gamma(u)$ , а  $\Gamma(u) = \Gamma(y)$ . Отсюда мы нашли  $w \in \Gamma(y) \cap \chi_1$ . Если  $\chi_1$  — парадигматическая категория, то доказательство ведется абсолютно так же с заменой  $\Gamma$  на  $S$  и  $S$  на  $\Gamma$ . Теорема доказана.

**Т. 6. 3.** Если язык  $S$ - $T$ -правилен, то для любого  $T$  множество всех  $T$ -диагностических фраз есть правильный фрагмент.

**Доказательство.** Пусть  $x \in T$ ,  $x \in S(y)$  и  $fxg$  есть  $T$ -диагностическая фраза. Поскольку  $x \in S(y)$ , то цепочка  $fxg$  есть фраза. Предположим, что она не является  $T$ -диагностической. Если  $(f, g)$  есть  $T$ -диагностический контекст, то фраза  $fxg$  также  $T$ -диагностическая, ибо в силу  $S$ - $T$ -правильности из  $x \in S(y) \cap T$  следует, что  $y \in T$ . Если же  $(f, g)$  не является  $T$ -диагностическим контекстом, то поскольку  $fxg$  есть  $T$ -диагностическая фраза, она имеет вид: а)  $fxg = h_1 x h_2 x h_3$ , где  $(h_1 x h_2, h_3)$  (1) есть  $T$ -диагностический контекст и  $z \in T$  или б)  $fxg = h_1 z h_2 x h_3$ , где  $(h_1, h_2 x h_3)$  (2) есть  $T$ -диагностический контекст и  $z \in T$ . Поскольку случаи а) и б) различаются лишь входждением  $z$  в  $f$  или  $g$ , а это для доказательства неважно, то мы ограничимся лишь первым случаем. Предположим, что контекст  $(h_1 y h_2, h_3)$  (3) не допускает парадигму  $\Gamma(w)$  такую, что  $\Gamma(w) \subseteq T$ . В силу  $T$ -диагностичности контекста (1) существует  $w' \in \Gamma(w)$  такое, что цепочка  $h_1 y h_2 w' h_3$  есть фраза, а стало быть контекст (3) допускает  $\Gamma(w)$ . Полученное противоречие доказывает теорему.

**Т. 6. 4.—6. 5.** См. доказательство теоремы 4. 4 в «Моделях языка».

**Т. 6. 6.** Определения 6. 18 и 6. 18' равносильны.

**Доказательство** (в доказательстве мы будем пользоваться термином «подкласс» для объекта, определенного в 6. 18 и «штрихованный подкласс» для объекта, определенного в 6. 18'). Пусть  $x$  и  $y$  принадлежат одному подклассу. Тогда существует парадигма  $\Gamma(w)$  такая, что  $PS(x) \cap \Gamma(w) \neq \emptyset$  и  $PS(y) \cap \Gamma(w) \neq \emptyset$ . Из определения 6. 18' следует, что каждая парадигма целиком содержится в штрихованном подклассе, а из определения 6. 17 следует, что каждое подсемейство целиком содержится в штрихованном подклассе. Поскольку  $PS(x) \cap \Gamma(w) \neq \emptyset$ , подсемейство  $PS(x)$  включено в тот же штрихованный подкласс, что и  $\Gamma(w)$ , а поскольку  $PS(y) \cap \Gamma(w) \neq \emptyset$ , то  $\Gamma(w)$  включено в тот же штрихованный подкласс, что и  $PS(y)$ , но  $PS(x)$  включено в тот же штрихованный подкласс, что и  $PS(y)$ , а стало быть  $x$  входит в тот же штрихованный подкласс, что и  $y$ . Пусть теперь наоборот  $x$  и  $y$  принадлежат одному штрихованному подклассу  $M'$ . Образует подкласс слова  $x$ , т. е. объединение всех подсемейств, пересекающихся с  $\Gamma(x)$ . Раз  $x$  и  $y$  принадлежат одному штрихованному подклассу,  $\Gamma(x) \cap S(y) \neq \emptyset$ . Возьмем произвольное  $x' \in \Gamma(x)$ . Из 6. 17 и 6. 18' следует, что пересечение штрихованного подкласса и семейства есть подсемейство. Поэтому из  $x' \in S(y)$  и из принадлежности  $x'$  и  $y$  к одному штрихованному подклассу следует, что  $x' \in PS(y)$ . Итак  $x' \in PS(y) \cap \Gamma(x)$ , а значит  $PS(y)$  входит в тот же подкласс, что и  $x$ , а стало быть, и  $y$  входит в тот же подкласс, что и  $x$ . Теорема доказана.

**Т. 6. 7.** Для того, чтобы два слова принадлежали к одному роду в смысле Маркуса, необходимо и достаточно, чтобы для любого слова  $x' \in \Gamma(x)$  нашлось такое слово  $y'$ , что  $y' \in \Gamma(y) \cap S(x')$  или для любого слова  $y' \in \Gamma(y)$  нашлось  $x$  такое, что  $x' \in \Gamma(x) \cap S(y')$ .

**Доказательство.** Пусть два слова  $x$  и  $y$  принадлежат к одному роду в смысле Маркуса. Возьмем произвольное  $x'' \in \Gamma(x)$  и произвольное  $y'' \in \Gamma(y)$ . Нам надо найти  $z \in \Gamma(x) \cap S(y'')$  или  $z \in \Gamma(y) \cap S(x'')$ . Но  $\Gamma$ - $S$ -цепь,

связывающая  $x''$  и  $y''$ , может, очевидно, представлять собой один из следующих четырех случаев:

1.  $w \in \Gamma(x'') \cap S(y'')$
2.  $w \in S(x'') \cap \Gamma(y'')$
3.  $w \in \Gamma(x'') \cap \Gamma(y'')$
4.  $w \in S(x'') \cap S(y'')$

В первом случае для произвольного  $y''$  нашлось  $w$  такое, что  $w \in \Gamma(x) \cap S(y'')$ , во втором случае для произвольного  $x''$  нашлось  $w \in \Gamma(y)$  такое, что  $w \in S(x'')$ . В третьем и четвертом случае тривиальным образом выполнены оба условия. Пусть теперь для любого  $x' \in \Gamma(x)$  найдется  $y' \in \Gamma(y)$  такое, что  $y' \in S(x')$ . Пусть  $z \in \Gamma(x)$  и  $w \in \Gamma(y)$ . Тогда по условию верно одно из двух:  $w \in \Gamma(x) \cap S(z)$  или  $w \in \Gamma(y) \cap S(z)$ . В обоих случаях  $\Gamma$ - $S$ -цепь состоит не более чем из трех слов. Теорема доказана.

**Т. 6. 8.** Подкласс есть синтагматическая категория тогда и только тогда, когда язык формально однороден.

**Доказательство.** Пусть  $x \in S(y)$  и язык однороден.

Предположим, что  $x$  и  $y$  входят в разные подклассы. Это означает, что существует  $x' \in \Gamma(x)$ , для которого в  $\Gamma(y)$  не найдется слова, входящего в то же семейство. Покажем, что этого не может быть. В самом деле  $x' \in \Gamma(x') \cap S(y)$ . В силу однородности существует  $z \in S(x') \cap \Gamma(y)$ . Итак, соответствующее слово найдено. Стало быть  $x$  и  $y$  относятся к одному подклассу, а значит подкласс — категория синтагматическая.

Теперь пусть, наоборот, любой подкласс есть синтагматическая категория и пусть  $z \in \Gamma(x) \cap S(y)$ . Пусть подкласс слова  $y$  есть синтагматическая категория, тогда  $z$  и  $y$  входят в один подкласс и в парадигме  $\Gamma(y)$  найдется слово  $w$ , входящее в одно семейство с  $x$ . Теорема доказана.

**Т. 6. 9.** Если все категории, не совпадающие с подклассом, синтагматические и язык формально однороден, то в нем два слова, относящиеся к одному семейству, относятся к одному категориальному пучку.

**Доказательство.** Пусть  $x \in S(y)$  и  $x \in x \cap \sigma$ , где  $x$  абстрактная категория рода, а  $\sigma$  пересечение нескольких синтагматических категорий. Допустим, что  $y$  входит в другой категориальный пучок. Но раз язык формально однороден, то в силу теоремы 6. 8,  $y$  входит в тот же подкласс. По условию других категорий, которые могли бы быть несинтагматическими, нет. Все синтагматические категории у обоих слов совпадают. Стало быть,  $x$  и  $y$  относятся к одному категориальному пучку.

**Т. 7. 1.** Не существует ограниченной грамматики, не порождающей ни одного пустого слова и совместимой с языком  $L^*$ .

**Доказательство.** Пусть существует такая грамматика.

Тогда первым ее правилом должно быть

$$C_1 \rightarrow c C_2.$$

Раз грамматика однозначна, то не может уже быть никакого перехода  $C_1 \rightarrow c C_3$ , а поскольку производятся фразы  $cb$  и  $ce$ , то должны также с необходимостью вводиться правила

$$C_2 \rightarrow b C_3$$

$$C_2 \rightarrow e C_4.$$

Наличие фраз  $cde$  и  $cab$  требует введения правил

$$C_2 \rightarrow a C_2 \quad \text{или} \quad C_2 \rightarrow a C_5$$

$$C_2 \rightarrow d C_2 \quad C_5 \rightarrow b C_6$$

$$C_2 \rightarrow d C_7$$

$$C_7 \rightarrow e C_8$$

Однако во втором случае грамматика была бы омонимичной.

В первом случае грамматика произведет все необходимые фразы, однако одновременно будет произведена запрещенная цепочка, например, *сае*. Мы перепробовали все возможности построения грамматики без использования пустых слов. Таким образом, теорема доказана.

**Т. 7. 2.** Язык, в котором имеется хотя бы одна фраза с рамкой  $(i, j)$ , несовместим с ограниченной грамматикой.

**Доказательство.** Пусть  $f = x_1 \dots x_{i-1} x_i x_{i+1} \dots x_{j-1} x_j x_{j+1} \dots x_n$  (1) есть фраза с рамкой  $(i, j)$ , производимая ограниченной грамматикой. Тогда фраза  $g = x_1 \dots x_{i-1} y_i x_{i+1} \dots x_{j-1} y_j x_{j+1} \dots x_n$  (2) также должна производиться этой грамматикой, в то время как цепочка  $h = x_1 \dots x_{i-1} y_i x_{i+1} x_{i+1} \dots x_{j-1} x_j x_{j+1} \dots x_n$  (3) запрещена и не должна производиться грамматикой. Рассмотрим ребро графа состояний, помеченное словом  $x_{i-1}$  (в силу ограниченности такое ребро только одно); пусть оно состоит из состояний  $(C_k, C_{k+1})$ . При переходе из состояния  $C_{k+1}$  должно порождаться два слова  $y_i$  и  $x_i$ , итак, пусть словом  $x_i$  помечено ребро  $(C_{k+1}, C_l)$ , а словом  $y_i$  ребро  $(C_{k+1}, C_{l+1})$ . Заметим, теперь, что словом  $x_{i+1}$  может быть помечено лишь одно ребро  $(C_{l+2}, C_{l+3})$ , и для того, чтобы производились фразы (1) и (2) необходимо пометить ребра  $(C_l, C_{l+1})$  и  $(C_l, C_{l+2})$  пустыми словами. Легко заметить, что какой бы путь по графу мы теперь ни выбрали, мы придем в состояние  $C_m$ , при выходе из которого выдается как  $x_j$  так и  $y_j$ , но тогда наша грамматика произведет и цепочку (3). Теорема доказана.

**Т. 7. 3—7. 5.** См. доказательства теорем 3—5: А. В. Гладкий. Конфигурационные характеристики языков. «Проблемы кибернетики», вып. 10. М., 1963, стр. 255—260.

**Т. 7. 6.** См.: И. И. Ревзин. Об одной синтаксической модели. — ВЯ, 1963, № 2, стр. 149.

## ИСПРАВЛЕНИЯ К КНИГЕ «МОДЕЛИ ЯЗЫКА»

Автор просит тех читателей, которые имеют в своем распоряжении экземпляры книги «Модели книги» (Изд-во АН СССР. М. 1962), внести в них следующие изменения и уточнения:

стр. 24, строка 6—7 снизу. В определении фонемы вычеркнуть слово «неоднородных».

стр. 27, 2 строка снизу. Пункт б) следует формулировать следующим образом: «между  $a_{i+1}$  и  $a_i$  можно установить взаимно-однозначное соответствие так, чтобы соответствующие признаки либо совпадали, либо были однородны».

стр. 28, 9 строка снизу: Пункт б) следует формулировать: «Процедура, описанная в пункте 5, повторяется до тех пор, пока не будут исчерпаны все архифонемы, общие хотя бы двум фонемам подсистемы».

стр. 34, 18—19 строка сверху. Ядро фонологической подсистемы следует определить следующим образом: «Ядром называется любая неустая, фонологически однородная подсистема. Некоторое разбиение допустимо для ядра, если любые две фонемы, входящие в один класс этого разбиения, либо обе принадлежат, либо обе не принадлежат данному ядру. Главным ядром подсистемы называется ядро, допускающее наибольшее число разбиений» (пример приведенный в книге, дает как раз главное ядро).

стр. 66, 16 строка снизу. Определение семейства дополнить по данной книге (см. 0. 3. 30).

стр. 69, 22—23 строка снизу. Вместо «элемент  $B_i$ » и «элемент  $B_j$ » нужно «класс  $B_i$ » и «класс  $B_j$ ».

стр. 71, 10 строка сверху. При интерпретации «окрестности» вместо слов «то для любого другого» следует написать: «для достаточно большого числа других».

стр. 86, 25 строка сверху. Нужно добавить: «на фрагменте, не содержащем прилагательных».

стр. 86, 10 строка снизу. Нужно добавить: «если оставаться в пределах форм 3-го лица».

стр. 91. Определение подсемейства уточнить в соответствии с 0.6.17 данной книги. Аналогично уточнить определение подокрестности (ввести симметричные требования).

стр. 93, конец примера вставить 13. *окна | стола*

14. *ламп | столов*

стр. 94, 19 строка снизу: вычеркнуть слово «неоднородных».

стр. 94, 13 строка снизу. Добавить: «или объединением подсемейств».

стр. 95, 20 строка сверху. Добавить: «если отвлечься от *лицо* (в смысле: *должностное лицо*), а также *чудовище*, *насекомое*, *животное*, *млекопитающее* и т. п.».

стр. 95, 4 строка снизу. Вычеркнуть: «неоднородные».

стр. 97, 9 строка сверху. Добавить: (если отвлечься от отдельных слов типа *щец*, *дровец*, *некто*, *нечто*, *полжизни*, *себя*).

стр. 102, 24—25 строка сверху. Пункт 2) следует исправить так: «для всякого слова  $y' \in S(x)$  и имеющего вид  $i_j d_i$ , найдется слово  $x' \in \Gamma(y')$  и имеющее вид  $i_j d_i$ ».

стр. 103. В примере, даваемом под  $A$  и  $B$  следует разграничить  $-oi_1$  (например, в *водяной*) и  $-oi_2$  (например, в *водой*) как нетождественные (и оговорить, что подобные случайные совпадения не нарушают общего принципа).

стр. 112, 1—2 снизу. В формулировке теоремы 4.7. Добавить условия, что  $S_i \neq S_j$ .

стр. 113, 1 строка сверху. Вместо  $S_j \rightarrow S_i$  нужно  $S_i \rightarrow S_j$ .

Пятой главой вообще не рекомендуется пользоваться, поэтому для нее мы не приводим исправлений.

## БИБЛИОГРАФИЯ

В Библиографию включены работы, непосредственно связанные с лингвистическим неколичественным моделированием. В нее не вошли работы, посвященные статистическому моделированию, а также моделированию лексикологии и семантики. Это объясняется как тем, что данные вопросы непосредственно не затрагиваются в основной части книги, так и наличием специальных библиографий по этим вопросам. По этой же причине исключена литература, специально посвященная вопросам машинного перевода и информационно-логических систем.

Приводимая Библиография лишь частично пересекается с цитированной в тексте книги литературой, в которой часто приводятся ссылки на литературу по общему и славянскому языкознанию, непосредственно не связанную с моделированием. Автор стремился сделать библиографию по возможности полной, чтобы она имела самостоятельную ценность. Однако поскольку ему не известна вся литература в данной области и, кроме того, ему приходилось производить отбор тех работ, которые представлялись ему существенными (по достигнутым результатам или общим идеям), то, естественно, данная Библиография не может претендовать на полноту, а скорее отражает тот круг работ, который послужил основой для написания книги, с одной стороны, и относится к узкой области неколичественного лингвистического моделирования, с другой.

Литература на русском языке включает и переводную, при этом в целях экономии места соответствующие работы уже не приводятся в списке иностранной литературы.

- А н д р е е в Н. Д. Моделирование языка на базе его статистической и теоретико-множественной структуры. «Тезисы совещания по математической лингвистике». Л., 1959, стр. 15—22.
- А н д р е е в Н. Д. Методы статистико-комбинаторного анализа языка в действии и в перспективе. «Статистико-комбинаторное моделирование языков». М., 1965, стр. 5—20.
- А н д р е е в Н. Д., З и н д е р Л. Р. Основные проблемы прикладной лингвистики. — ВЯ, 1959, № 4, стр. 3—19.
- А р с е н т ь е в а Н. Г. О синтезе предложений русского языка при помощи машины. «Научно-техническая информация», 1963, № 6, стр. 24—31; № 7, стр. 21—29.
- А р с е н т ь е в а Н. Г. О синтезе предложений русского языка при помощи машины. «Проблемы кибернетики», вып. 10. М., Физматгиз, 1963.
- А р с е н т ь е в а Н. Г. О двух способах порождения предложений русского языка. «Проблемы кибернетики», вып. 14. М., изд-во «Наука», 1965.
- А р с е н т ь е в а Н. Г. Исследование текстов, синтезированных машиной. «Проблемы кибернетики», вып. 15. М., изд-во «Наука», 1965.

- А х м а н о в а О. С., М е л ь ч у к И. А., П а д у ч е в а Е. В., Ф р у м к и н а Р. М. О точных методах исследования языка. М., 1961.
- Б а б и ц к и й К. И. К вопросу о моделировании структуры простого предложения. «Проблемы структурной лингвистики». М., 1962, стр. 115—129.
- Б а р - Х и л л е л И. О рекурсивных определениях в эмпирических науках. «Математическая лингвистика» (сб. переводов). М., 1964, стр. 101—107.
- Б а р - Х и л л е л И. Разрешающие процедуры для структуры естественных языков, там же, стр. 108—121.
- Б а р - Х и л л е л И. Некоторые новые результаты в теоретической лингвистике. «Математическая логика и ее применения». Перев. с англ. М., 1965, стр. 273—280.
- Б е л е ц к и й М. И. Несколько моделей языка. «Доклады на конференции по обработке информации, машинному переводу и автоматическому чтению текста». М., 1961.
- Б е л е ц к и й М. И. Модель русского языка, описывающая простые предложения без однородности. «Научно-техническая информация», 1964, № 7, стр. 37—42.
- Б е л е ц к и й М. И., Г р и г о р я н В. М., З а с л а в с к и й И. Д. Аксиоматическое описание порядка и управления слов в некоторых типах предложений. «Математические вопросы кибернетики и вычислительной техники», № 1. Ереван, 1963, стр. 71—85.
- Б е л о з е р о в В. Н. Формальное определение фонемы. — ВЯ, 1964, № 6, стр. 54—60.
- Б о р о д и н В. В. К модели описания языка. «Доклады на конференции по обработке информации, машинному переводу и автоматическому чтению текста», вып. 6. М., 1961.
- Б р а т ч и к о в И. Л. Некоторые проблемы формальной морфологии, там же, вып. 2.
- Б р а т ч и к о в И. Л., Ф и т и а л о в С. Я., Ц е й т и н Г. С. О структуре информации для машинного перевода, там же, вып. 2.
- Б у р л а к о в а М. И., Н и к о л а е в а Т. М., С е г а л Д. М., Т о п о р о в В. Н. Структурная типология и славянское языкознание. «Структурно-типологические исследования». М., 1962, стр. 3—18.
- В а к у л о в с к а я Г. В., К у л а г и н а О. С. Об одном способе анализа текста. «Проблемы кибернетики», вып. 12. М., изд-во «Наука», 1964.
- В о л о ц к а я З. М. Установление отношения производности между словами (опыт применения трансформационного метода). — ВЯ, 1960, № 3, стр. 100—107.
- В о л о ц к а я З. М., М о л о ш н а я Т. Н., Н и к о л а е в а Т. М., Р е в з и н И. И., Ц и в и ш н а Т. В. Об одном подходе к типологии славянских языков (на материале системы склонения существительного). «Славянское языкознание. Доклады советской делегации». М., 1963, стр. 510—552.
- В о л о ц к а я З. М., М о л о ш н а я Т. Н., Н и к о л а е в а Т. М. Опыт описания русского языка в его письменной форме. М., 1965.
- Г л а д к и й А. В. Конфигурационные характеристики языков. «Проблемы кибернетики», вып. 10. М., 1963, стр. 251—260.
- Г л а д к и й А. В. Об одном способе формализации понятия синтаксической связи. «Проблемы кибернетики», вып. 11. М., 1964.
- Г л а д к и й А. В. Алгоритм распознавания конфигураций для класса автоматных языков. «Проблемы кибернетики», вып. 12. М., 1964, стр. 243—245.
- Г л а д к и й А. В. Исследования по теории порождающих грамматик. (Автореф. докт. дисс.). Новосибирск, 1965.
- Г л а д к и й А. В. О распознавании замещаемости в рекурсивных языках. «Алгебра и логика. Семинар», т. 2, вып. 3. Новосибирск, 1963, стр. 5—22.
- Г л а д к и й А. В. Грамматики с линейной памятью. «Алгебра и логика. Семинар», т. 2, вып. 5. Новосибирск, 1963, стр. 43—55.

- Г л а д к и й А. В. Алгоритмическая природа инвариантных свойств грамматик непосредственно составляющих. «Алгебра и логика. Семинар», т. 3, вып. 2. Новосибирск, 1964, стр. 17—32.
- Г л а д к и й А. В. О сложности вывода в грамматиках непосредственно составляющих. «Алгебра и логика. Семинар», т. 3, вып. 5—6. Новосибирск, 1964, стр. 29—44.
- Г л а д к и й А. В. Некоторые алгоритмические проблемы для контекстно свободных грамматик. «Алгебра и логика. Семинар», Новосибирск, т. 4, вып. 1, 1965, стр. 3—13.
- Г л а д к и й А. В. Алгоритмическая нераспознаваемость существенной неопределенности контекстно-свободных языков. «Алгебра и логика. Семинар», т. 4, вып. 4. Новосибирск, 1965, стр. 53—64.
- Г л а д к и й А. В. Прямое доказательство теоремы Мэтьюза, там же, стр. 65—70.
- Г л а д к и й А. В. О формальных методах в лингвистике. — ВЯ, 1966, № 3.
- Г л а д к и й А. В. Лекции по математической лингвистике для студентов НГУ. Новосибирск, 1966.
- Д а н е ш Фр. Опыт теоретической интерпретации синтаксической омонимии. — ВЯ, 1964, № 6, стр. 3—16.
- Д о б р у ш и н Р. Л. Элементарная грамматическая категория. «Бюллетень объединения по проблемам машинного перевода», № 5. М., 1957, стр. 19—21.
- Д о б р у ш и н Р. Л. Математические методы в лингвистике. «Математическое просвещение», 1961, № 6, стр. 37—60.
- Д р е й з и н Ф. А. Некоторые свойства связей между словами в предложении. «Научные труды Ташкентского гос. ун-та. Серия Математика», вып. 228, стр. 33—39. Ташкент, 1963.
- З а л и з н я к А. А. Лингвистические задачи. «Исследования по структурной типологии». М., 1963, стр. 137—159.
- З а л и з н я к А. А. «Условное ударение» в русском склонении. — ВЯ, 1964, № 1, стр. 14—29.
- З а л и з н я к А. А. Категории рода и одушевленности в русском языке. — ВЯ, 1964, № 4.
- З а л и з н я к А. А. Об использовании понятий «автоматической выводимости» и «зависимого признака» при описании знаковых систем. «Симпозиум по структурному изучению знаковых систем. Тезисы докладов». М., 1962, стр. 55.
- З а л и з н я к А. А. О возможной связи между операционными понятиями синхронного описания и диахронией, там же, стр. 56.
- З а л и з н я к А. А. Классификация и синтез именных парадигм современного русского языка. (Автореф. канд. дисс.). М., 1965.
- З а л и з н я к А. А., П а д у ч е в а Е. В. О связи языка описаний с родным языком. «Программа и тезисы докладов в летней школе по вторичным моделирующим системам». Тарту, 1964, стр. 7—9.
- З а л и з н я к А. А., И в а н о в Вяч. Вс., Т о п о р о в В. Н. О возможности структурно-типологического изучения некоторых моделирующих систем. «Структурно-типологические исследования». М., 1962, стр. 134—143.
- З а с о р и н а Л. Н. Замечания о математической лингвистике. «Вестник МГУ», серия VII, 1962, № 1, стр. 55—59.
- З а с о р и н а Л. Н. Трансформации как метод лингвистического эксперимента. «Трансформационный метод в структурной лингвистике». М., 1964.
- З в е г и н ц е в В. А. Применение в лингвистике логико-математических методов. «Новое в лингвистике», вып. IV. М., 1965, стр. 7—32.
- З и н о в ъ е в А. А. О математической лингвистике. «Вопросы философии», 1959, № 9, стр. 133—140.
- З и н о в ъ е в А. А., Р е в з и н И. И. Логическая модель как средство научного исследования. «Вопросы философии», 1960, № 1, стр. 82—90.
- И в а н о в Вяч. Вс. Код и сообщение. «Бюллетень объединения по проблемам машинного перевода», 1957, № 5, стр. 48—50.



- Иванов Вяч. Вс. *л-Мерное пространство языка*, там же, стр. 51—53.
- Иванов Вяч. Вс. *Языкознание и математика*. «Бюллетень объединения по проблемам машинного перевода», 1957, № 3, стр. 5—10.
- Иванов Вяч. Вс. *О построении информационного языка для текстов по дескриптивной лингвистике*. «Доклады на конференции по обработке информации, машинному переводу и автоматическому чтению текста», вып. 7. М., 1961.
- Иванов Вяч. Вс. *Язык в сопоставлении с другими средствами передачи и хранения информации*, там же (см. сокращенный текст: «Прикладная лингвистика и машинный перевод». Киев, 1962, стр. 79—113).
- Иванов Вяч. Вс. *Теорема Геделя и лингвистические парадоксы*. «Тезисы конференции по машинному переводу». М., 1958, стр. 42—43.
- Иванов Вяч. Вс. *О языках с максимальной гибкостью и языках с минимальной гибкостью*. «Тезисы докладов на V Международном методическом семинаре преподавателей русского языка высших учебных заведений социалистических стран». М., 1962, стр. 21—22.
- Иванов Вяч. Вс. *Некоторые проблемы современной лингвистики*. «Народы Азии и Африки», 1963, стр. 156—178.
- Иванов Вяч. Вс., Топоров В. Н. *К реконструкции праславянского текста*. «Славянское языкознание. Доклады советской делегации». М., 1963, стр. 88—157.
- Ингве В. *Гипотеза глубины*. Перев. с англ. «Новое в лингвистике», вып. IV, стр. 126—138.
- Иорданская Л. Н. *Свойства правильной синтаксической структуры и алгоритм ее обнаружения (на материале русского языка)*. «Проблемы кибернетики», вып. 11. М., 1964.
- Иорданская Л. Н. *О некоторых свойствах правильной синтаксической структуры*. — ВЯ, 1963, № 4.
- Исаченко А. В. *Трансформационный анализ кратких и полных прилагательных*. «Исследования по структурной типологии». М., 1963, стр. 61—93.
- Карпинская О. Г. *Типология рода в славянских языках*. — ВЯ, 1964, № 6, стр. 61—76.
- Карпинская О. Г. *К сопоставлению родовых систем славянских и германских языков*. «Тезисы докладов на II Всесоюзной конференции по славяно-германскому языкознанию». Минск, 1965, стр. 39—43.
- Карпинская О. Г. *Методы типологического описания славянских родовых систем*. «Лингвистические исследования по общей и славянской типологии». М., 1966.
- Карри Г. *Некоторые логические аспекты грамматической структуры*. Перев. с англ. «Новое в лингвистике», вып. IV. М., 1965, стр. 98—116.
- Кузнецов П. С. *Об основных положениях фонологии*. — ВЯ, 1959, № 2.
- Кузнецов П. С. *Опыт формального определения слова*. — ВЯ, 1964, № 3.
- Кулагина О. С. *Об одном способе определения грамматических понятий на базе теории множеств*. «Проблемы кибернетики», вып. 1. М., 1958, стр. 203—214.
- Кулагина О. С. *Составление при помощи машины алгоритма анализа текста*. «Доклады на конференции по обработке информации, машинному переводу и автоматическому чтению текста», вып. 9. М., 1961.
- Курилович Е. *Понятие изоморфизма*. Перев. с франц. «Очерки по лингвистике». М., 1962, стр. 21—36.
- Ламбек И. *Математическое исследование структуры предложений*. «Математическая лингвистика» (сб. переводов). М., 1964, стр. 47—68.
- Левин Ю. И. *Математика и язык*. «Математика в школе», 1962, № 5, стр. 5—11.
- Левин Ю. И. *Об описании системы лингвистических объектов, обладающих общими свойствами*. — ВЯ, 1964, № 4, стр. 112—119.
- Левин Ю. И. *Знаки, язык, математика*. «О некоторых вопросах современной математики и кибернетики». М., 1965, стр. 428—461.

- Лейкина Б. М. Некоторые аспекты характеристики валентностей. «Доклады на конференции по обработке информации, машинному переводу и автоматическому чтению текста», вып. 5. М., 1961.
- Лекомцев Ю. К. Элементы теории языковой сочетаемости. «Проблемы структурной лингвистики. 1963». М., 1963, стр. 23—46.
- Лекомцева М. И., Сегал Д. М., Судник Т. М., Шур С. М. Опыт построения фонологической типологии близкородственных языков. «Славянское языкознание. Доклады советской делегации». М., 1963, стр. 423—476.
- Ленский Д. Н. Об одном применении алгебры бинарных отношений в лингвистике. «Уч. зап. Кавказского университета», вып. 16, Серия физико-математическая. Нальчик, 1962.
- Лесерф И. Применение программы и модели конфликтной ситуации к автоматическому синтаксическому анализу естественных языков «Научно-техническая информация», 1963, № 10, стр. 42—50.
- Лиз Р. Б. О возможностях проверки лингвистических положений. — ВЯ, 1962, № 4, стр. 47—55.
- Ломковская М. В. Исчисление, порождающее ядерные русские предложения, ч. I. «Научно-техническая информация», 1965, № 7, стр. 35—41; ч. II, там же, 1965, № 9, стр. 37—40.
- Линдбах Я. Принципы философского языка. Опыт точного языкознания. Пг., 1915.
- Лосев А. Ф. О трудностях изложения математической лингвистики для лингвистов. — ВЯ, 1965, № 5.
- Маркус С. Об одной логической модели элементарной грамматической категории, III. «Revue roumaine de mathématiques pures et appliquées», vol. 7, 1962, № 4, стр. 683—691.
- Маркус С. Теория графов, лингвистические оппозиции и инвариантная структура. «Проблемы структурной лингвистики». М., 1962, стр. 22—30.
- Маркус С. Логический аспект лингвистических оппозиций, «Проблемы структурной лингвистики. 1963». М., 1963, стр. 47—74.
- Маркус С. Грамматический род и его логическая модель. «Математическая лингвистика» (сб. переводов). М., 1964, стр. 122—144.
- Мартемьянов Ю. С. К построению языка лингвистических описаний. «Симпозиум по структурному изучению знаковых систем. Тезисы докладов». М., 1962, стр. 10—12.
- Мельчук И. А. К вопросу о грамматическом в языке-посреднике. «Машинный перевод и прикладная лингвистика». М., 1960, № 4, стр. 25—45.
- Мельчук И. А. О терминах «устойчивость» и «идиоматичность». — ВЯ, 1960, № 4, стр. 73—80.
- Мельчук И. А. К вопросу о термине «система» в лингвистике. «Zeichen und System der Sprache», II. Berlin, 1962.
- Мельчук И. А. О «внутренней флексии» в индоевропейских и семитских языках. — ВЯ, 1963, № 4, стр. 27—40.
- Мельчук И. А. Автоматический синтаксический анализ. Новосибирск, 1964.
- Молошная Т. Н. О понятии грамматической конфигурации. «Структурно-типологические исследования». М., 1962, стр. 46—59.
- Мэтьюз Д. Г. Разрывность и асимметрия в грамматиках непосредственно составляющих. «Математическая лингвистика» (сб. переводов). М., 1964, стр. 150—159.
- Небеский Л. Об одной формализации разбора предложения, там же, стр. 145—149.
- Небеский Л. К одной модели анализа предложения. «The Prague Bulletin of mathematical linguistics», 1964, № 2, стр. 25—45.
- Новотный М. Об алгебраизации теоретико-множественной модели языка, «Проблемы кибернетики», вып. 15. М. 1965, стр. 234—243.

- Падучева Е. В. О способах представления синтаксической структуры предложения. — ВЯ, 1964, № 2, стр. 99—113.
- Падучева Е. В. Некоторые вопросы перевода с информационно-логического языка на русский «Научно-техническая информация», 1964, № 2, стр. 20—27.
- Падучева Е. В. Синтез сложных предложений с однозначной синтаксической структурой (при переводе с информационно-логического языка на русский), там же, 1964, № 6, стр. 43—49.
- Падучева Е. В. О понятии конфигурации. — ВЯ, 1965, № 1, стр. 56—68.
- Падучева Е. В. О связях глубины по Ингве со структурой дерева подчинения, «Научно-техническая информация». № 6, 1966, стр. 38—43.
- Паненова Я., Сгал П., Математическая и структурная лингвистика в ЧССР, «Эзик и литература», № 6, 1964, стр. 1—8.
- Плат У. Математическая лингвистика. Перев. с англ. «Новое в лингвистике», вып. 4. М., 1965, стр. 201—245.
- Путнам Х. Некоторые спорные вопросы теории грамматики. Перев. с англ. «Новое в лингвистике», вып. 4. М., 1965, стр. 66—96.
- Ревзин И. И. О логической форме лингвистических определений. «Применение логики в науке и технике». М., 1960, стр. 140—148.
- Ревзин И. И. Формальный и семантический анализ синтаксических связей в языке, там же, стр. 128—139.
- Ревзин И. И. О некоторых понятиях так называемой теоретико-множественной концепции языка. — ВЯ, 1960, № 6, стр. 88—94.
- Ревзин И. И. О понятиях однородного языка и языка с полной трансформацией. «Структурно-типологические исследования». М., 1962, стр. 19—24.
- Ревзин И. И. О некоторых вопросах дистрибутивного анализа и его дальнейшей формализации (применительно к синтаксическому анализу). «Проблемы структурной лингвистики». М., 1962, стр. 13—21.
- Ревзин И. И. Модели языка. М., 1962.
- Ревзин И. И. Некоторые замечания о методах введения математических терминов в лингвистику. «Славянская лингвистическая терминология». София, 1963.
- Ревзин И. И. К типологии форм степеней сравнения в славянских языках. «Исследования по структурной типологии». М., 1963, стр. 35—41.
- Ревзин И. И. Некоторые формальные особенности парадигмы глагола (к проблеме внутренней типологии), там же, стр. 94—102.
- Ревзин И. И. Об одной синтаксической модели. — ВЯ, 1963, № 2, стр. 148—150.
- Ревзин И. И. К логическому обоснованию теории фонологических признаков. — ВЯ, № 5, 1964.
- Ревзин И. И. Некоторые вопросы теории моделей языка. «Научно-техническая информация», 1964, № 8, стр. 42—46.
- Ревзин И. И. Лингвистическое моделирование и обучение языку. — РЯНШ, 1964, № 6.
- Ревзин И. И. От структурной лингвистики к семиотике. «Вопросы философии», 1964, № 9.
- Ревзин И. И. Распространение теоретико-множественной модели на случай грамматической омонимии. «Научно-техническая информация», 1965, № 3.
- Ревзин И. И. Структурная лингвистика и единство языкознания. — ВЯ, 1965, № 3.
- Ревзин И. И., Розенцвейг В. Ю. Основы общего и машинного перевода. М., 1965.
- Реформатский А. А. Дихотомическая классификация дифференциальных признаков и фонематическая модель языка. «Вопросы теории языка в современной зарубежной лингвистике». М., 1961, стр. 106—122.

- Реформатский А. А. О перекодировании и трансформации коммуникативных систем. «Исследования по структурной типологии». М., 1963, стр. 208—215.
- Соболева П. А. Опыт исчисления трансформаций на основе теории С. К. Шаумяна о порождении классов слов в процессе порождения грамматики. «Проблемы структурной лингвистики». М., 1963., стр. 233—263.
- Сухотин Б. В. Алгоритмы лингвистической дешифровки, «Проблемы структурной лингвистики». 1963». М., 1963.
- Тоноров В. Н. О трансформационном методе. «Трансформационный метод в структурной лингвистике». М., 1964, стр. 74—87.
- Трубецкой Н. С. Основы фонологии. Перев. с нем. М., 1960.
- Уитфильд Ф. Критерии для модели языка. «Математическая логика и ее применения». М., 1965, стр. 307—313.
- Уорт Д. С. Об отображении линейных отношений в порождающих моделях. — ВЯ, 1964, № 5.
- Успенский Б. А. Типологическая классификация языков как основа языковых соответствий. — ВЯ, № 6, 1961.
- Успенский Б. А. Принципы структурной типологии. М., 1962.
- Успенский Б. А. Опыт трансформационного исследования синтаксической типологии. «Исследования по структурной типологии». М., 1963, стр. 52—60.
- Успенский Б. А. Структурная типология языков. М., 1965.
- Успенский В. А. К определению падежа по Колмогорову. «Бюллетень объединения по проблемам машинного перевода», 1957, № 5, стр. 11—18.
- Успенский В. А. К определению части речи в теоретико-множественной системе языка, там же, стр. 22—26.
- Успенский В. А. Одна модель для понятия фонемы. — ВЯ, 1964, № 6, стр. 39—53.
- Фитиалов С. Я. Формально-математические модели языков и структура алгоритмов перевода. «Тезисы совещания по математической лингвистике». Л., 1959.
- Фитиалов С. Я. О построении формальной морфологии в связи с машинным переводом. «Доклады на конференции по обработке информации, машинному переводу и автоматическому чтению текста», вып. 2. М., 1961.
- Фитиалов С. Я. О моделировании синтаксиса в структурной лингвистике. «Проблемы структурной лингвистики». М., 1962, стр. 100—114.
- Фитиалов С. Я. О двух типах исчислений. «Научно-техническая информация», 1964, № 7, стр. 30—36.
- Фитиалов С. Я. Трансформация в аксиоматических грамматиках. «Трансформационный метод в структурной лингвистике». М., 1964 стр. 3, 2, 11.
- Халле М. О роли простоты в лингвистических описаниях. Перев. с англ. «Новое в лингвистике», вып. IV, 1965, стр. 117—125.
- Хауген Э. Направления в современном языкознании. «Новое в лингвистике», вып. 1. М., 1960, стр. 249—251.
- Херари Ф., Пейпер Г. К построению общего исчисления распределения фонем. «Математическая лингвистика» (сб. переводов). М., 1964, стр. 11—46.
- Хилл А. О грамматической отмеченности предложений. Перев. с англ. — ВЯ, 1962, № 4.
- Холодович А. А. Опыт теории подклассов слов. — ВЯ, 1960, № 1.
- Хомский Н. Три модели описания языка. «Кибернетический сборник» (сб. переводов), № 2. М., 1961, стр. 237—266.
- Хомский Н. О некоторых формальных свойствах грамматик. «Кибернетический сборник», № 5. М., 1962, стр. 279—311.
- Хомский Н. Заметка о грамматиках непосредственных составляющих, там же, стр. 312—315.

- Хомский Н. Синтаксические структуры. «Новое в лингвистике», вып. 2. М., 1962.
- Хомский Н. Несколько методологических замечаний о порождающих грамматиках. Перев. с англ. — ВЯ, 1962, № 4.
- Хомский Н. О понятии «правило грамматики». Перев. с англ. «Новое в лингвистике», вып. IV. М., 1965, стр. 34—65.
- Хомский Н. Логические основы лингвистической теории. Перев. с англ. «Новое в лингвистике», вып. IV. М., 1965, стр. 465—475.
- Хомский Н. Объяснительные модели в лингвистике. «Математическая логика и ее применения». М., 1965, стр. 245—272.
- Хомский Н. Лингвистика, логика, психология и вычислительные устройства, «Математическая лингвистика» (сб. переводов). М., 1964, стр. 69—100.
- Хомский Н., Миллер Д. Языки с конечным числом состояний. «Кибернетический сборник» (сб. переводов), № 4, стр. 231—255.
- Хэррис З. С. Совместная встречаемость и трансформации в языковой структуре. «Новое в лингвистике», вып. II. М., 1962.
- Цейтин Г. С. К вопросу о построении математических моделей языка. «Доклады на конференции по обработке информации, машинному переводу и автоматическому чтению текста», вып. 3. М., 1961.
- Цейтин Г. С., Засорина Л. Н. О выделении конфигураций в русском предложении, там же.
- Цивьян Т. В. Имя существительное в балканских языках (к структурно-типологической характеристике балканского союза). М., 1965.
- Чжао Юань-жень. Модели в лингвистике и модели вообще. «Математическая логика и ее применения». Перев. с англ. М., 1965, стр. 281—292.
- Шаумян С. К. История системы дифференциальных элементов в польском языке. М., 1958.
- Шаумян С. К. Логический анализ понятия структуры языка. «Тезисы конференции по машинному переводу». М., 1958, стр. 65—66.
- Шаумян С. К. Логический анализ понятия фонемы. «Логические исследования». М., 1959, стр. 159—177.
- Шаумян С. К. Генерализация и постулирование конструкторов в изучении структуры языка. «Тезисы совещания по математической лингвистике». Л., 1959, стр. 38—43.
- Шаумян С. К. Операционные определения и их применение в фонологии. «Применение логики в науке и технике». М., 1960.
- Шаумян С. К. Двухступенчатая теория фонемы и дифференциальных элементов. — ВЯ, 1960, № 5, стр. 18—34.
- Шаумян С. К. Проблемы теоретической фонологии. М., 1962.
- Шаумян С. К. Преобразование информации в процессе познания и двухступенчатая теория структурной лингвистики. «Проблемы структурной лингвистики». М., 1962, стр. 5—11.
- Шаумян С. К. Панхроническая система дифференциальных элементов и двухступенчатая теория фонологии, там же, стр. 75—79.
- Шаумян С. К. Порождающая лингвистическая модель на базе принципа двухступенчатости. — ВЯ, 1963, № 2, стр. 57—71.
- Шаумян С. К. О логическом базисе лингвистической теории. «Проблемы структурной лингвистики. 1963». М., 1963, стр. 3—8.
- Шаумян С. К. Структурная лингвистика. М., 1965.
- Шаумян С. К., Соболева П. А. Аппликативная порождающая модель и исчисление трансформаций в русском языке. М., 1963.
- Шрейдер Ю. А. Свойства проективности языка. «Научно-техническая информация», 1964, № 8.
- Шрейдер Ю. А. Характеристики сложности структуры текста. «Научно-техническая информация», 1966, 17, стр. 34—38.
- Якобсон Р., Фант Г. М., Халле М. Введение в анализ речи. Перев. с англ. «Новое в лингвистике», вып. II. М., 1962.

- A b e r n a t h y R. The problem of linguistic equivalence. PSAM: [Proceedings of Symposia in applied mathematics], vol. 12. «Structure of language and its mathematical aspects», ed. by R. Jakobson, Providence, 1961, crp. 86—95.
- A b r a h a m S. A formal study of generative grammars (1). «Computational Linguistics», II. Budapest, 1964, crp. 5—18.
- A b r a h a m S. Some Questions of Phrase Structure Grammars, 1. «Computational Linguistics», IV. Budapest, 1965, crp. 61—70.
- A j d u k i e w i c z K. Die syntaktische Konnexität. «Studia Philosophica», 1. Lemberg, 1935, crp. 1—27.
- A n d r e y e v N. D. Models as a tool in development of linguistic theory. «Word», 1962, № 1—2, crp. 186—197.
- A p o s t e l L. Towards the formal study of models in nonformal sciences. «Synthese», 1960, № 12.
- A p o s t e l L., M a n d e l b r o t B., M o r f A. Logique, langue, théorie de l'information. Paris, 1957.
- B a c h E. An Introduction to Transformational Grammars. New York, 1964.
- B a r - H i l l e l Y. Four lectures on computational linguistics and machine translation. Jerusalem, 1963.
- B a r - H i l l e l Y. A quasi-arithmetical notation for syntactic description. «Language», № 29, 1953, crp. 47—58.
- B a r - H i l l e l Y. On syntactical categories. «Journal of symbolic Logic», 1950, № 15.
- B a r - H i l l e l Y. Three methodological remarks on «Fundamentals of Language». — «Word», 1957, № 13, crp. 323—335.
- B a r - H i l l e l Y., G a i f m a n C. and S h a m i r E. On categorial and phrase-structure grammars. «Bulletin of the Research Council of Israel», F 9. Jerusalem, 1960, crp. 1—16.
- B a r - H i l l e l Y. and S h a m i r F. Finite-state languages: Formal representation and adequacy problems, там же, F 8, 1960, crp. 155—166.
- B a r - H i l l e l Y., P e r l e s M., S h a m i r E. On formal properties of simple phrase structure grammars. «Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung», 1961, № 14, crp. 143—172.
- B a t ó g T. Logiczna rekonstrukcja pojęcia fonemu. «Studia logica», vol. XI, 1961, crp. 139—183.
- B a z e l l C. E. Syntactic Relations and Linguistic Typology. «Cahiers F.de Saussure», 8, 1949, crp. 5—20.
- B e l e v i t c h V. Propositions pour une théorie mathématique de la structure du langage. «Revue de questions scientifiques», vol. 16, 1955.
- B e l e v i t c h V. Langage des machines et langage humain. Paris, 1956.
- B e l l e r t I. Relational Phrase Structure Grammar and its Tentative Applications. «Information and Control», Vol. 8, 1965, № 5, crp. 503—529.
- B e n z e c r i J. P. Physique et Langue. «La Traduction Automatique», 1963, № 2, crp. 31—50.
- B e r k a K. O nektých aplikacích formální logiky v jazykovědě. «Slovo a slovesnost», 1961, № 3, crp. 198—203.
- B e r k a K., N o v a k P. «Výklad fonologických a gramatických pojmů pomocí pojmů teorie množin (nad knihou I. I. Revzina «Modeli jazyka») «Slovo a slovesnost», 1963, № 2.
- B r o d d a B., K a r l g r e n H. Relative positions of elements in linguistic strings. «Statistical Methods in linguistics», 1964, № 3, crp. 49—101.
- B i e r w i s c h M. Über den theoretischen Status des Morphems. «Studia grammatica», № 1. Berlin, 1962.
- B i e r w i s c h M. Grammatik des deutschen Verbs. «Studia grammatica», № 2. Berlin, 1963.
- B r a f f o r t P. Éléments de linguistique mathématique. «Euratom. Enseignement préparatoire aux techniques de la documentation automatique». Bruxelles, 1960, crp. 51—85.

- Camion P. Analyse algébrique élémentaire du critère de Lecerf-Ihm. «Rapport 9 RISA, Euratom», № 3, 1960, стр. 1—7.
- Cantineau J. Les oppositions significatives. «Cahiers F. de Saussure», vol. 10, 1952, стр. 1—10.
- Cantineau J. Le classement logique des oppositions. «Word», vol. 11, 1955, № 1, стр. 1—9.
- Carnap R. Logical Syntax of Language. New York, 1937.
- Carroll J. B. An operational model for language behavior. «Anthropological Linguistics», 1, 1959, № 1, стр. 37—54.
- Ceccato S. Operational linguistics and translation. «Methodos», № 12, 1960, стр. 11—80.
- Yuen Ren Chao. The non-uniqueness of phonemic solutions of phonetic systems. «Readings in linguistics». New York, 1958, стр. 38—54.
- Chatman S. Immediate Constituents and Expansion Analysis. «Word», 1955, vol. 11, № 3.
- Cherry E. C. Roman Jakobsons «distinctive features» as the normal coordinates of a language. «For Roman Jakobson». The Hague, 1956.
- Cherry E. C., Halle M., Jakobson R. Toward the logical description of languages in their phonemic aspects. «Language», 1953, № 1.
- Chomsky N. Systems of syntactical analysis. «Journal of Symbolic Logic», vol. 18, 1953, № 3, стр. 242—256.
- Chomsky N. Logical Syntax and Semantics, Their linguistic relevance. «Language», 1955, № 1, стр. 36—45.
- Chomsky N. Logical structure in language. «American Documentation», vol. 8, 1957, стр. 284—291.
- Chomsky N., Schutzenberger M. P. The algebraic Theory of Context-free languages. «Computer Programming and formal systems. Studies in logic», Amsterdam, 1963.
- Chomsky N., Miller G. Introduction to the formal analyses of natural languages. «Handbook of Mathematical Psychology». New York, 1963.
- Craciun C. Sur la notion de racine dans la théorie algébrique de la grammaire. «Revue roumaine de mathématiques pures et appliquées», № 2, 1964.
- Craciun C. Sur une généralisation de la catégorie grammaticale au sens de Dobrushin, «Cahiers de linguistique théorique et appliquée», № 2, 1964.
- Cooper W. S. Set Theory and Syntactic Description. London — the Hague — Paris, 1964.
- Čulík K. On Languages Generated by Some Types of Algorithms. «Prague Studies in Mathematical Linguistics». 1 Praha, 1966, стр. 141—146.
- Čulík K. Some notes on finite state languages. «Časopis pro pestování matematiky», № 86, 1961, стр. 43—55.
- Curry H. B. Some logical aspects of grammatical structure. — PSAM, vol. 12, стр. 56—68.
- Daneš F. A three-level approach to syntax. «Travaux linguistiques de Prague, 1, l'école de Prague, d'aujourd'hui». Prague, 1964, стр. 225—239.
- Delavenay E. Vers une linguistique générale répondant aux besoins de la traduction automatique. «La Traduction automatique», 1961, № 2, стр. 71.
- Diderichsen L. The importance of distribution versus other criteria in linguistic analysis. «Reports for 8-th International Congress of linguists», vol. 1, 1957.
- Fodor J. La typologie des langues slaves et le genre grammaticale. «Славянска филология», т. III. София, 1963, стр. 51—58.
- Gaifman C. Dependency systems and phrase-structure systems. «Rand Corporation», Santa Monica, California, 1961.

- Gårding L. Relations and order. «Studia linguistica», 6, 1955, crp. 21—34.
- Garvin P. L. From model to procedure. «Proceedings of the National Symposium on machine translation». London, 1961, crp. 367—370.
- Garvin P. L. Automatic linguistic analysis — a heuristic problem. «Proceedings of the first International conference on machine translation of languages and applied language analysis». Teddington, 1962.
- Goodman N. Graphs for linguistics. — PSAM, vol. 12, 1961, crp. 51—55.
- Greenberg J. Essays in linguistics. New York, 1957.
- Greibach S. Formal Systems in linguistics. RAND Corporation. Santa Monica. California, 1962.
- Halle M. The strategy of phonemics. «Word», vol. 10. «Linguistics today», 1954, crp. 197—209.
- Halle M. In the defence of the number two. «Studies presented to J. Whatmough». 's-Gravenhage, 1957.
- Halle M. Phonology in generative grammar. «Word», vol. 18, 1962, № 1—2, crp. 54—71.
- Harris Z. S. Structural linguistics. Chicago, 1961.
- Harris Z. S. The Transformational model of Language Structure. «Anthropological Linguistics», I, 1959, № 1, crp. 27—29.
- Harris Z. S. String analysis of sentence structure. The Haag, 1962.
- Harwood F. W. Axiomatic syntax, The construction and the evaluation of a syntactic calculus. «Language», 1955, № 3, crp. 409—413.
- Hays D. G. Grouping and dependency theories, The RAND Corporation P—1910. Santa Monica, California, 1960.
- Hays D. G. Dependency theory: a formalism and some observations. «Language», 1964, № 4, crp. 510—514.
- Herzberger H. The joints of English. — PSAM, vol. 12, crp. 93—103.
- Hirschberg L., Lynch I. Discussions sur l'hypothèse de projectivité. «Rapport Cetus, Euratom», 35, 1961.
- Hiž H. Questions and Answers. «The Journal of Philosophy», vol. LIX, 1962, № 10.
- Hiž H. The Role of Paraphrase in Grammar. «Report of the 15-th Annual RTM on Linguistic and Language Studies», № 17, 1964.
- Hiž H. Congrammaticality, batteries of transformations and grammatical categories. — PSAM, vol. 12, crp. 43—50.
- Hockett C. F. Two models of grammatical description. «Word», 1954, № 2—3, crp. 210—234.
- Horecki J. Morfematička štruktura slovenčiny. Bratislava, 1964.
- Isačenko A. Grammatičnost a význam. «Acta Universitatis Carolinae. Philologica, Slavica pragensia», IV, 1962, crp. 47—51.
- Jakobson R. Linguistics and communication theory. — PSAM, vol. 12. Providence, 1961.
- Jakobson R., Halle M. Fundamentals of language. 's-Gravenhage, 1956.
- Jakobson R. Shifters, Verbal Categories and the Russian verb. Harward, 1957.
- Jaurisová A., Jauris M. Užití theorie množin v jazykovědě. «Slovo a slovesnost», 1960, № 21, crp. 43—41.
- Jos M. Description of language design. «Journal of the Acoustical society of America», 1950, № 22, crp. 701—708.
- Kanger St. The notion of a phoneme. «Statistical methods in linguistics», 1964, № 3, crp. 43—48.
- Van Katwijk A. A grammar of Dutch number names. «Foundations of language», № 1, 1965, crp. 51—58.
- Katz J. J., Postal P. M. An Integral Theory of linguistic Description. Cambridge (Mass), 1964.



- Kiefer F. Some aspects of mathematical models in linguistics. «Statistical methods in linguistics», 1964, № 3, crp. 8—26.
- Kiefer F. Some Questions of Semantic Theory. «Computational Linguistics», N 4. Budapest, 1965, crp. 71—78.
- Kiefer F., Abraham S. How Computers understand Sentences. «Computational linguistics», III. Budapest, 1964, crp. 1—28.
- Koschmieder E. Die Mathematisierung der Sprachwissenschaft. «Forschungen und Fortschritte», 39 Jg., Juli 1956, Heft 7.
- Lecerf Y. Une représentation algébrique de la structure des phrases dans diverses langues naturelles. «Comptes rendus de l'Académie des Sciences», vol. 232. Paris, 1961, № 2.
- Lecerf Y., Ihm P. Eléments pour une grammaire générale des langues projectives. «Euratom, Rapport Cetus», 1960, № 1.
- Lees R. B. The grammar of English nominalizations. «International Journal of American linguistics», 26, 1960, № 3, crp. 11.
- Longacre R. E. Grammar Discovery procedures. A field manual. The Haag, 1964.
- Marcus S. Unele semnificații ale lingvisticii matematice. «Limba română», vol 3, 1960, crp. 3—14.
- Marcus S. Structures linguistiques et structures topologiques. «Revue roumaine de mathématiques pures et appliquées», vol. 6, 1961, № 3, crp. 501—506.
- Marcus S. Description, à l'aide de la théorie des ensembles, de certains phénomènes morphologiques, *там же*, vol. 6, 1961, No 4, crp. 735—744.
- Marcus S. Noțiunea de algoritm și implicațiile ei lingvistice. «Limba română», 1961, № 3, crp. 204—211.
- Marcus S. Unele aspecte ale lingvisticii matematice. in Uniunea Sovietică. «Analele romino-sovietice, seria matematică-fizică», 1961, № 3, crp. 3—34; № 4, crp. 3—34.
- Marcus S. Sur un modèle logique de la catégorie grammaticale élémentaire, I. «Revue roumaine de mathématiques pures et appliquées», vol. 7, 1962, № 1, crp. 91—107.
- Marcus S. Sur un modèle logique de la catégorie grammaticale élémentaire, II. «Zeitschrift für mathematische Logik und Grundlagen der Mathematik», Bd., 8, 1962, № 3—4, crp. 323—329.
- Marcus S. Asupra unui model logic al părții de vorbire. «Studii și cercetări matematice», vol. 13, 1962, № 1, crp. 37—62.
- Marcus S. Un criteriu contextual de clasificare a cuvintelor (cu aplicație la adjectivalele din limba română). «Studii și cercetări lingvistice», vol. 13, 1962, № 2, crp. 177—189.
- Marcus S. A synchronic analysis of the grammatical gender. «Revue de linguistique», vol. 7, 1963, № 1, crp. 99—111.
- Marcus S. Automates finis, progressions arithmétiques et grammaires à un nombre fini d'états. «Comptes rendus de l'Académie des sciences de Paris», 256, 1963, № 17, crp. 3571—3574.
- Marcus S. Lingvistică matematică. București, 1963.
- Marcus S. Langues complément adéquates et langues régulières. «Zeitschrift für mathematische Logik und Grundlagen der Mathematik», vol. 10, 1964, № 1.
- Marcus S. Modelarea matematică a limbii. «Limba română», vol. 5, 1963, crp. 478—492.
- Marcus S. Typologie des langues et modèles logiques «Acta Mathematica Academiae Scientiarum Hungaricae», vol. 14, 1963, № 3—4, crp. 269—281.
- Marcus S. Aspecte ale modelării matematice în lingvistică. «Studii și cercetări lingvistice», vol. XIX, 1963, crp. 487—501.
- Marcus S. Un model matematic al fonemului. «Studii și cercetări matematice», vol. XIV, 1963, № 3, crp. 405—421.

- Marcus S. Modèles mathématiques pour la catégorie grammaticale du cas. «Revue roumaine de mathématiques pures et appliquées», vol. 8, 1963, № 4, crp. 585—610.
- Marcus S. Sur la notion de la projectivité. «Computational linguistics», III. Budapest, 1964, crp. 75—92.
- Marcus S. Grammatici și automate finite. București, 1964.
- Marcus S., Vasiliu Em. Mathématiques et phonologie. Théorie des graphes et consonantisme de la langue roumaine. «Revue roumaine de mathématiques pures et appliquées», vol. 5, 1960, № 2, crp. 319—340; № 3—4, crp. 681—704.
- Mathews G. H. Analysis by synthesis of natural languages. «Proceedings of the 1961 International Conference on Machine Translation of languages and Applied language analysis», vol. II. London, 1962, crp. 532—540.
- Nebeský L. The  $\xi$ -Grammar. «Prague Studies in Mathematical Linguistics», 1. Praha, 1966, crp. 147—154.
- Nebeský L. O jednom algebraickem modelu jazyka, «Slovo a slovesnost», 1963, № 4, crp. 231—237.
- Nebeský L. Conditional Replacement of Words. «The Prague Bulletin of Mathematical Linguistics», N 3. Praha, 1965, crp. 3—12.
- Nebeský L. K pojmu věty smysluplné a věty gramaticky správné. «Kibernetika a její využívání». Praha, 1965.
- Nebeský L., Sgall P. Vztah formy a funkce v jazyce (Pokus o axiomatizaci). «Slovo a slovesnost», 1962, № 3, crp. 174—189.
- Nebeský L., Sgall P. O relacích a operacích v jazyce. «Slovo a slovesnost», 1965, № 2.
- Novák P. Nekteré otázky syntaktické analýzy (z hlediska strojového překladu). «Slovo a slovesnost», 1962, № 1, crp. 9—20.
- Novák P. K vytváření terminologie matematické lingvistiky. «Československý terminologický časopis». Bratislava, crp. 234—237.
- Novák P. On Mathematical Models of Linguistic Objects. «Prague Studies in Mathematical Linguistics», 1. Praha, 1966, crp. 155—157.
- Novotný M. Matematika pro lingvistu. Praha, 1965.
- Novotný M. On some Algebraic Concepts of Mathematical Linguistics. «Prague Studies in Mathematical Linguistics». Praha, 1966, crp. 125—140.
- Oettinger A. G. Linguistics and mathematics. «Studies presented to J. Whatmough», 1957, crp. 179—186.
- Parker-Rhodes A. F. Is there an interlingual element in syntax. «Proceedings of 9-th International Congress of linguists». The Haag, 1964, crp. 176—190.
- Parker-Rhodes A. F. A new Model of Syntactic Description. «1961 International Conference on Machine Translation and Applied Language Analysis», vol. 1. London, 1962.
- Postal P. M. Nový vývoj teorii transformační gramatiky. «Slovo a slovesnost», 1965, № 1, crp. 1—13.
- Ray R. S. The logic of linguistics. «Methodos», vol. 13, 1962, crp. 51—52, 239—254.
- Sgall P. K užívání matematických pojmů v lingvistice. «Slovo a slovesnost», 1965, № 1, crp. 83—85.
- Sgall P. Generativní systémy v lingvistice. «Slovo a slovesnost», 1964, № 4, crp. 274—282.
- Sgall P. Soustava padových koncovek v češtině. «Acta Universitatis Carolinae. Philologica», 1960, № 2, crp. 65—84.
- Sgall P. Zur Frage der Ebenen im Sprachsystem. «Travaux linguistiques de Prague, L'école de Prague d'aujourd'hui», № 1. Prague, 1964, crp. 95—109.
- Sgall P., Novák P. Matematické metody a marxistická jazykověda. «Problémy marxistické jazykovědy». Praha, 1962, crp. 41—50.

- Sgall P. a kolektiv. Cesty moderní jazykovědy. Jazykověda a automatizace. «Mala moderní Encyklopedie 47». Orbis Praha, 1964.
- Shamir E. On sequential languages and two classes of regular events. «Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung». Bd. 18, Heft. 1. 1965, стр. 61—70.
- Shamir E. A remark on discovery algorithms for grammars. «Information and controls», vol. 5, 1962, стр. 246—251.
- Spang-Hansen H. On the simplicity of description. «Travaux du Cercle linguistique de Copenhague», vol. 5, 1949.
- Spang-Hansen H. Probability and structural classification in language description. Copenhagen, 1959.
- Spang-Hansen H. Mathematical linguistics — a trend in name or in fact? «Proceedings of 9th International Congress of linguists». The Hague, 1964, стр. 61—71.
- Šaumjan S. K. Concerning the logical basis of linguistic theory, там же, стр. 155—160.
- Tesnière L. Éléments de syntaxe structurale. Paris, 1959.
- Ungeheuer G. Das logische Fundament binärer Phonemklassifikationen. «Studia Linguistica», 1959, стр. 69—97.
- Varga D. Yngve's Hypothesis and some problems of the Mechanical Analysis. «Computational linguistics», № 3. Budapest, 1964, стр. 47—77.
- Vasiliu Em. Une classification des consonnes roumaines d'après le critère de la distribution. «Mélanges linguistiques». Bucarest, 1957, стр. 97—112.
- Vasiliu Em. Contribution à la théorie du signe linguistique. «Cahiers de linguistique théorique et appliquée», vol. 1, 1962, стр. 245—253.
- Yngve V. H. A model and a hypothesis for language structure. «Proceedings of the American Philosophical society», vol. 104, 1960, № 4, стр. 444—466.
- Zelinka B. Un langage adéquat non homogène, dont les classes sont disjointes deux à deux. «Revue roumaine de mathématiques pures et appliquées», t. X, 1965, № 8, стр. 1255—1258.

## УКАЗАТЕЛЬ ТЕРМИНОВ

Приводятся как термины, получившие определение в тексте книги — в этом случае указан только номер определения (первая цифра соответствует номеру главы, а вторая — номеру определения), — так и термины, а также лингвистические явления, не определяемые в книге, — в этом случае слово набрано курсивом и указываются по возможности все параграфы, в которых обсуждается данное понятие.

- Абстракция отождествления* § 1  
— *потенциальной осуществимости* § 7  
*агглютинативность* § 32, 34, 37  
автоматная грамматика см. грамматика с конечным числом состояний 3. 13  
алгоритм 2.29  
*алфавитное разбиение* § 21  
аморфное слово 2.2  
аморфный язык 5.1  
*аналогия* § 19, 34  
антидиагностический фрагмент 6.14  
антисимметричность 2.12  
*артикуль* § 32, 49  
атрибут квазиконфигурации 7.21  
атрибут конфигурации 7.19
- База оппозиции** 2.40  
бесконтекстная грамматика 3.12
- Валентность** 7.41  
вершина графа 2.15  
взаимно-однозначное отображение 2.27  
взаимозамещаемость 3.26  
*вид глагола* § 34, 35  
вполне правильное слово 5.11  
— правильный язык 5.11  
— простое слово 5.14  
— простой язык 5.14  
вхождение 3.2  
вывод 3.9
- Генус* § 45, 46  
гиперсемейство 6.22  
гипертип 4.5  
синтаксически обусловленный г. 4.11
- главное слово составляющей 7.24  
глагол § 29, 37, 43, 46  
глубина 7.9  
гомоморфизм 2.26  
грамматика 3.8  
— без омонимии 7.3  
— с конечным числом состояний 3.13  
— непосредственно составляющих 3.11  
*грамматическая категория* § 15  
граф 2.15
- Дезигнат* § 1  
*денотат* § 1  
дерево 2.22  
— логических возможностей 2.28  
*дескриптивный подход* § 1  
диагностическая фраза 6.9  
диагностический контекст 6.8  
— фрагмент 6.10  
диаграмма грамматики 7.1  
дизъюнктивная оппозиция 2.42  
дистрибутивная близость 3.4  
доминанционная фраза и доминанционный язык 7.32  
доминанция 7.31  
дополнение множества 2.6  
дополнительное отношение 2.10  
— распределение 2.37
- Единичное разбиение* § 21  
емкость контекста 3.31
- Залоговые формы глагола* § 58  
замещение 3.25  
*замкнутая предметная область* § 1  
*замкнутый язык* § 17

запрещенная цепочка § 17  
знак § 1

Идеально простое слово 5.15  
— простой язык 5.15  
избирательность 7.42  
изолированная оппозиция 2.44  
изоморфизм 2.27  
индекс транспозиции 7.36  
исправимая фраза § 17

**Базис** § 43  
категориально однородный язык 6.6  
категориальный пучок 3.1  
категоризация § 1, 2  
категория § 15  
— состояния § 27  
квазиконфигурация 7.20  
класс разбиения 2.7  
— связности 3.46  
— управления § 43, 46  
коммутация 2.36  
конечно-характеризуемый язык 7.16  
конструктивный подход § 1  
контекст 3.20  
контекстный класс 3.33  
конфигурация 7.12—7.13  
конфигурационный язык 7.22  
корень дерева 2.22  
кортеж § 16  
куст дерева 7.23

Левовициклический символ 7.7  
линейная упорядоченность 2.23  
лицо глагола § 44

Местоимение § 22, 26, 28, 34, 37,  
42, 43, 46  
метрическое пространство 2.33  
многосторонняя оппозиция 2.42  
множество § 7  
модель § 4  
морфемная структура слова § 49

**Наречие** § 27, 31, 41, 47  
начальный класс 3.39  
непосредственная выводимость 3.9  
— связанность 3.45  
непосредственное подчинение § 18  
непосредственно составляющая 3.10  
неприводимая фраза 7.15

Область определения и область значения 2.25  
образ 2.25  
— цепочки 3.17

обоюдный род § 34, 47  
обратно-алфавитное разбиение § 21  
объединение  
— множеств 2.4  
— отношений 2.17  
объемлющее множество 2.1  
объемности (принцип) § 6  
ограниченная грамматика 7.3  
ограниченное замещение 3.36  
ограниченно однородное слово 5.23  
— однородный язык 5.21  
однозначная грамматика 7.2  
означаемое и означающее § 1  
омонимия § 15  
— синтаксическая § 18  
омонимичное слово 3.42  
— первого рода 3.43  
— второго рода 3.44  
операционное определение § 11  
оппозиция § 13  
определитель 3.6  
основной тезис теории моделей § 22  
отглагольное существительное § 36,  
37  
отмеченный образ 3.18  
категориальный пучок 6.5  
отношение 2.9  
отображение 2.24

Падеж 6.7—6.7  
падежеобразующий класс 6.11  
парадигматика § 1  
парадигма § 14, 35  
парадигматическая категория 6.2  
парадигматически полный язык 5.17  
пересечение 2.5  
перечислимое множество 2.32  
подграф 2.16  
подкласс 6.18  
подмножество 2.1  
подрод 6.25  
подсемейство 6.17  
подчинение синтаксическое 3.3  
подфраза 7.28  
подязык 3.38  
полнозначное слово § 14, 30  
полностью однородное слово 5.22  
— однородный язык 5.20  
порождающий процесс 2.31  
правильное слово 5.10  
— укрупнение 2.8  
правильный фрагмент 3.35  
— язык 5.9  
правоциклический символ 7.7  
предельно грамматикализованное слово 5.6  
— грамматикализованный язык 5.5  
— служебное слово 5.8

привативная оппозиция 2.39  
*принадлежности (отношение)* § 7  
*прилагательное* § 27, 28, 31, 43, 46  
примыкание 7.43  
производное разбиение 3.29  
производность (слова) 5.30  
прообраз 2.25  
пропорциональная оппозиция 2.43  
простая конфигурация 7.14  
простое слово 5.12  
простой язык 5.12  
проективная фраза 7.37—7.38  
пустая клетка 2.28  
*пустое множество* § 7  
*пустая цепочка* § 16

Равенство (множеств) 2.2  
равномерно распределенная категория 6.3—6.4  
разбиение 2.7  
раздел 4.3  
разложимая конфигурация 7.18  
разрешимое множество 2.30  
ранг  
— конфигурации 7.13  
— подмножества 2.28  
— сегмента 7.25  
— фразы 7.25  
расстояние 2.33—2.34  
*реальная фраза* § 17  
ребро графа 2.15  
регулярная конфигурация 7.19  
регулярный сегмент 3.6  
рефлексивность 2.11  
род 6.20  
— по Зализняку 6.23  
— по Маркусу 6.19  
ряд, порожденный начальным классом 3.40

Самовставляющийся символ 7.8  
свободное варьирование 2.38  
связный граф 2.20  
сегмент 3.5  
*семантически правильная фраза* § 60  
*семантические признаки* § 61  
семантическое подчинение 7.47  
семейство 3.30  
*семиотика* § 1  
симметризация 2.19  
симметричность 2.12  
сильная проективность 7.37  
*синтагматика* § 1  
синтагматическая категория 6.1  
синтаксическая группа 7.3  
синтаксически родственные слова 7.17

система составляющих 3.10  
*системность* § 10  
слабая проективность 7.38  
слабое согласование 7.47  
*слово* § 14  
*словообразование* § 62  
*сложноподчиненное предложение* § 51  
служебное слово 5.7  
*смысл* § 1, 18  
собственно аморфное слово 5.3  
собственная категория падежа 6.12  
собственное подмножество 2.3  
совершенный образ 3.19  
совместимость грамматики с языком 3.15  
согласование 7.45  
согласование в сильном смысле 7.46  
согласовательный класс 6.21  
составляющая фразы 3.10  
*степени сравнения* § 31, 34  
*стилистика* § 19  
строгое включение 2.3  
структура 2.27  
*структурная лингвистика* § 3  
структурная типология языков 1.1  
субкатегория 6.15  
*субстантивация прилагательных* § 27  
*существительное* § 27, 28, 37

Тип 4.1  
— синтаксически обусловленный 4.9  
*типология* § 5  
*T-обусловленная парадигма* 5.27  
*T-ряд* 4.6  
транзитивность (отношения) 2.13  
транзитивное замыкание 2.18  
транспозиция 7.35  
трансформируемость 5.26  
трансформационное правило 7.10

Укрупнение (разбиения) 2.8  
управление 7.44  
упрощение (фразы) 7.30  
*уровни отмеченности* § 19  
устраняемая рамка 7.5

Фонологическое слово 2.35  
формально однородное слово 5.19  
— однородный язык 5.18  
*фраза* § 17  
— порождаемая грамматикой 3.10  
— с полной трансформацией 5.28  
— с рамкой 7.4  
— нулевого ранга 7.25  
функциональное отношение (см. отбражение)

Цепочка 3.2  
Г-S-цепь 4.2  
цикл графа 2.21  
циклический символ 7.6

**Числительное § 11**

число существительного § 15, 35,  
41, 45

Эквивалентное отношение 2.14  
эквивалентная оппозиция 2.41

элементарная грамматическая кате-  
гория 3.41

Ядро квазиконфигурации 7.21  
конфигурации 7.19  
сегмента 3.6

язык 3.37

— с ненулевой парадигматикой 5.16  
— с полной трансформацией 5.29  
— с равномерно суживающимися  
классами 7.27  
— с хорошо заданными парадигма-  
ми 5.25

## СПИСОК СОКРАЩЕНИЙ

В	— виный падеж	МО	— мужской одушевленный род
Д	— дательный падеж	наст. вр.	— настоящее время
ед. ч.	— единственное число	НЛМО	— нелично-мужской одушевленный род
Ж	— женский род	неперех.	— непереходный глагол
ЖНО	— женский неодушевленный род	несов.	— несовершенный вид
ЖО	— женский одушевленный род	П	— предложный падеж
Зв	— звательный падеж	перех.	— переходный глагол
И	— именительный падеж	положит.	— положительная степень прилагательного
Л	— лицо	Р	— родительный падеж
ЛМО	— лично-мужской одушевленный род	С	— средний род
М(ест.)	— местный падеж	СНО	— средний неодушевленный род
мп. ч.	— множественное число	СО	— средний одушевленный род
М	— мужской род	сов.	— совершенный вид
МНО	— мужской неодушевленный род	Т	— творительный падеж

## ОСНОВНЫЕ ОБОЗНАЧЕНИЯ

$G(x)$	— парадигма слова $x$	О	— отношение линейной упорядоченности
$D(x)$	— класс связности слова $x$	$P, Q, R$	— произвольные отношения
$E$	— единичное разбиение	$Ra(x)$	— раздел слова $x$
$M$	— множество	$S(x)$	— семейство слова $x$
$\Pi(x)$	— категориальный пучок, содержащий слово $x$	$T(x)$	— тип слова $x$
С	— состояние (на схеме языка с конечным числом состояний)	$U$	— (универсальное) множество
$B$	— произвольное разбиение	$V$	— множество словоформ, «словарь»
$C$	— падеж	$V_1$	— множество символов грамматических категорий
$D$	— отношение непосредственного (синтаксического) подчинения	$\alpha$	— множество реальных фраз
$F$	— произвольное отображение	$\beta$	— множество запрещенных цепочек
$G$	— грамматика	$\Delta$	— правильный фрагмент языка
$H(x)$	— гипертип слова $x$	$\delta$	— диагностический фрагмент языка
$J$	— индекс	$\varkappa$	— категория
$K$	— контекстный класс	$\Lambda$	— пустая цепочка
$L$	— язык		



- |          |  |             |   |
|----------|--|-------------|---|
| $\pi$    | — признак                              | $\in$       | — знак принадлежности элеме-<br>мента к множеству |
| $\Phi$   | — множество фраз                       | $\cup$      | — знак объединения                                |
| $\tau$   | — синтаксически обусловлен-<br>ный тип | $\cap$      | — знак пересечения                                |
| $\omega$ | — обратно-алфавитное раз-<br>биение    | $\emptyset$ | — пустое множество                                |

# СОДЕРЖАНИЕ

От автора . . . . .	3
<b>ГЛАВА ПЕРВАЯ.</b>	
<b>Принципы лингвистического моделирования . . . . .</b>	<b>9</b>
§ 1. Некоторые семиотические понятия и общие принципы анализа знаковых систем . . . . .	9
§ 2. О роли лингвистических категорий . . . . .	15
§ 3. О соотношении традиционной лингвистики и теории моделей . . . . .	16
§ 4. Понятие модели. Три типа лингвистических моделей . . . . .	23
§ 5. Теория моделей и структурная типология . . . . .	29
<b>ГЛАВА ВТОРАЯ.</b>	
<b>Множества и отношения. Их использование в лингвистике . . . . .</b>	<b>33</b>
§ 6. Вводные замечания . . . . .	33
§ 7. Понятие множества. Операции над множествами . . . . .	35
§ 8. Понятие бинарного отношения . . . . .	37
§ 9. Понятие отображения . . . . .	42
§ 10. Дерево логических возможностей и понятие алгоритма . . . . .	44
§ 11. Порождающий процесс и операционные определения . . . . .	50
§ 12. Понятия расстояния и признакового пространства . . . . .	52
§ 13. Предварительные лингвистические примеры . . . . .	54
<b>ГЛАВА ТРЕТЬЯ.</b>	
<b>Основные лингвистические понятия и операции . . . . .</b>	<b>60</b>
§ 14. Слово и парадигма . . . . .	60
§ 15. Грамматические категории . . . . .	61
§ 16. Понятия цепочки и вхождения . . . . .	64
§ 17. Реальные фразы и запрещенные цепочки . . . . .	65
§ 18. Синтаксическая зависимость и смысл фразы . . . . .	73
§ 19. Один способ порождения фраз на основе реальных и запрещенных цепочек . . . . .	77
§ 20. Понятия грамматики и множества фраз, порождаемых грамматикой . . . . .	81
§ 21. Разбиения словаря и образы цепочек . . . . .	85
§ 22. Понятия контекста и замещения. Семейство . . . . .	87

§ 23. Понятие правильного фрагмента . . . . .	91
§ 24. Язык и подязык . . . . .	92
§ 25. Элементарная грамматическая категория и проблема омонимии . . . . .	93
<b>ГЛАВА ЧЕТВЕРТАЯ.</b>	
<b>Моделирование части речи . . . . .</b>	<b>97</b>
§ 26. Понятие типа . . . . .	97
§ 27. Понятия гипертита и <i>T</i> -ряда . . . . .	100
§ 28. Использование идеи синтаксической обусловленности при моделировании части речи . . . . .	108
§ 29. Часть речи «глагол». Некоторые общие замечания о моделировании части речи . . . . .	110
<b>ГЛАВА ПЯТАЯ.</b>	
<b>Теоретико-множественная типология языков . . . . .</b>	<b>113</b>
§ 30. Понятия аморфного языка и предельно-грамматикализованного языка. . . . .	113
§ 31. Понятие <i>S-T</i> -правильного языка . . . . .	115
§ 32. Простой язык и нормализованный язык . . . . .	117
§ 33. Понятие полного языка . . . . .	119
§ 34. Понятие однородного языка . . . . .	119
§ 35. Вопрос о разграничении индивидуальных парадигм. Язык с хорошо заданными парадигмами . . . . .	127
§ 36. Отношение трансформируемости . . . . .	129
§ 37. Язык с <i>T</i> -обусловленными парадигмами и язык с полной трансформацией . . . . .	130
§ 38. Общая оценка теоретико-множественной типологии . . . . .	136
<b>ГЛАВА ШЕСТАЯ.</b>	
<b>Моделирование грамматических категорий славянских языков . . . . .</b>	<b>137</b>
§ 39. Некоторые предварительные определения . . . . .	137
§ 40. Модели функции падежа существительных. Падежи русского языка . . . . .	139
§ 41. Субкатегории. Вопрос о категориях числа существительного . . . . .	146
§ 42. Падежи в других славянских языках . . . . .	148
§ 43. Абстрактный падеж как основная непарадигматическая категория языка . . . . .	153
§ 44. Исследование категорий глагольного лица и рода в славянских языках . . . . .	155
§ 45. Модели функции рода . . . . .	158
§ 46. Абстрактный род как основная парадигматическая категория языка . . . . .	163
§ 47. Инвентарь категорий рода в славянских языках . . . . .	166
§ 48. Расстояние между категориями как типологический показатель . . . . .	169
<b>ГЛАВА СЕДЬМАЯ.</b>	
<b>Некоторые вопросы, связанные с синтагматическими моделями . . . . .</b>	<b>178</b>
§ 49. Лингвистический аспект грамматик с конечным числом состояний . . . . .	178

§ 50. Бесконтекстные грамматики . . . . .	192
§ 51. Некоторые замечания о трансформационной модели . . . . .	199
§ 52. Конфигурационная модель . . . . .	201
§ 53. Система составляющих и дерево синтаксического подчинения . . . . .	207
§ 54. Об одном способе восстановления категорий на базе реальных и запрещенных цепочек . . . . .	209
§ 55. Восстановление отношений синтаксического подчинения. Синтаксические группы . . . . .	214
§ 56. Понятие транспозиции и принцип проективности . . . . .	225
§ 57. Валентности . . . . .	234
§ 58. Согласование, управление, примыкание . . . . .	239
<b>ГЛАВА ВОСЬМАЯ.</b>	
<b>О возможности моделирования семантики . . . . .</b>	<b>248</b>
§ 59. Некоторые трудности построения семантических моделей . . . . .	248
§ 60. О понятии семантически правильной фразы . . . . .	249
§ 61. Семантика грамматических категорий . . . . .	253
§ 62. Об одном возможном способе формирования семантических категорий . . . . .	255
§ 63. Парадигма как семантическое слово . . . . .	266
§ 64. Уровни языка и типы моделей . . . . .	266
<b>П р и л о ж е н и я . . . . .</b>	<b>271</b>
Доказательства или ссылки на доказательства теорем . . . . .	271
Исправления к книге «Модели языка» . . . . .	275
<b>Библиография . . . . .</b>	<b>277</b>
<b>Указатель терминов . . . . .</b>	<b>291</b>
<b>Список сокращений . . . . .</b>	<b>295</b>
<b>Основные обозначения . . . . .</b>	<b>295</b>

*Исаак Иосифович Ревзин*

**Метод моделирования  
и типология славянских языков**

*Утверждено к печати  
Институтом славяноведения  
АН СССР*

Редактор издательства *Г. Н. Корозо*  
Художник *Э. Л. Эрман*  
Технический редактор *Р. М. Денисова*

Сдано в набор 11/X 1966 г.  
Подписано к печати 7/II 1967 г.  
Формат 60×90<sup>1/16</sup>. Бумага типографская № 2.  
Уч.-изд. л. 19,3. Тираж 3400 экз.  
Тип. зак. 1275.  
Цена 1 р. 34 к.

Издательство «Наука».  
Москва, К-62, Подсосенский пер., 21

---

1-я типография издательства «Наука».  
Ленинград, В-34, 9 линия, д. 12

ОПЕЧАТКИ И ИСПРАВЛЕНИЯ

Стр.	Строка	Напечатано	Должно быть
24	16 сл.	Как было указано в § 2	Как было указано в § 3
42	19 св.	а) антирефлексивно и антисимметрично	а) есть дерево
50	2 сл.	<i>девятьдесят</i>	<i>девяльдесять</i>
51	2 св.	<i>десять</i>	<i>десять</i>
142	9 »	$x \in T_1$	$x \in T_1$
160	2 »	singularia pluralia	singularia и pluralia
172	7 св.	(Д/М) = 7/8	з (Д/М) = 7/8
174	4 св.	стр. 176	стр. 173
271	13 сл.	а $fug \in \bar{\Delta}$	а $fug \in \bar{\Delta}$
284	3 »	17	№ 11
Оборот титула		369—66	390—66

И. И. Ревзин. Метод моделирования и типология славянских языков.



ИЗДАТЕЛЬСТВО «НАУКА»

\* МЕТОД МОДЕЛИРОВАНИЯ И ТИПОЛОГИЯ СЛАВЯНСКИХ ЯЗЫКОВ \*