4
П-78

АКАДЕМИЯ НАУК СССР

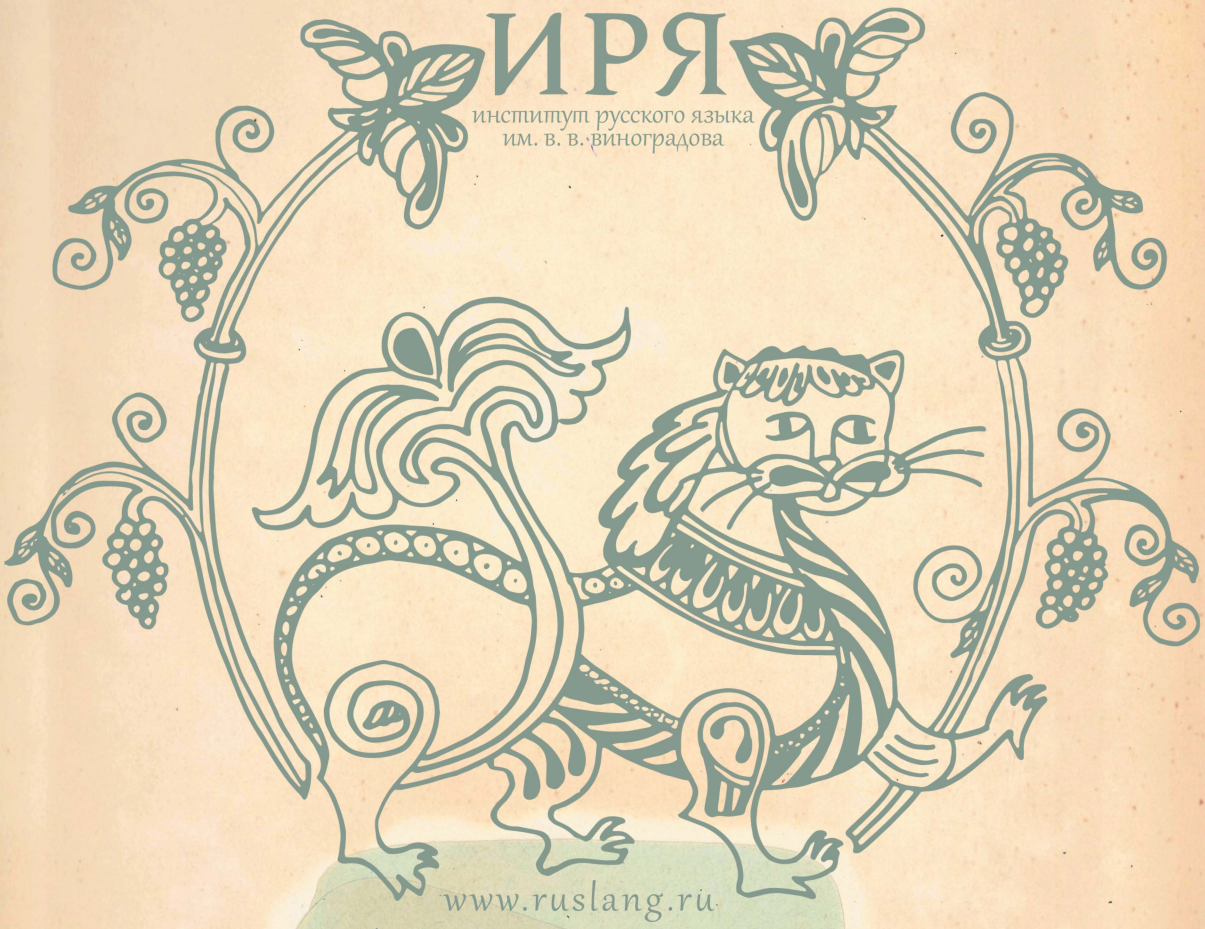
ПРОБЛЕМЫ СТРУКТУРНОЙ ЛИНГВИСТИКИ

1963

ИЗДАТЕЛЬСТВО АКАДЕМИИ НАУК СССР

ОЯ-отд. (старый)

ИРЯ
институт русского языка
им. в. в. виноградова



www.ruslang.ru

58967

АКАДЕМИЯ НАУК СССР
ИНСТИТУТ РУССКОГО ЯЗЫКА

ПРОБЛЕМЫ СТРУКТУРНОЙ ЛИНГВИСТИКИ

1963

ИЗДАТЕЛЬСТВО АКАДЕМИИ НАУК СССР
Москва 1963

Ответственный редактор

доктор филологических наук **С. К. ШАУМЯН**

4
П-78



58967

I. ОБЩИЕ ВОПРОСЫ

С. К. Шаумян

О ЛОГИЧЕСКОМ БАЗИСЕ ЛИНГВИСТИЧЕСКОЙ ТЕОРИИ

В качестве логического базиса лингвистической теории мы считаем необходимым принять принцип двухступенчатости, служащий стержнем современной логики науки.

Согласно этому принципу во всякой теоретической науке необходимо строго различать две ступени абстракции: ступень наблюдения и ступень конструктов. На ступени наблюдения мы имеем дело с прямо наблюдаемыми объектами, свойствами и отношениями, которые принято называть элементарными. К конструктам относятся недоступные прямому наблюдению объекты, свойства и отношения. Конструкты связываются со ступенью наблюдения посредством так называемых правил корреспонденции.

По поводу термина «ступень наблюдения» следует сделать следующее замечание. Ступень наблюдения надо понимать не в психологическом смысле, а в логическом смысле, т. е. как совокупность исходных фактов науки, подлежащих теоретическому объяснению. Вопрос о том, что наблюдается и что не наблюдается нами, всегда может вызвать споры, если подходить к нему под психологическим углом зрения. Но в логическом плане понятие ступени наблюдения абсолютно недвусмысленно. Речь идет именно о тех фактах, которые в рамках данной научной дисциплины определяются как наблюдаемые и которые по определению служат эмпирической базой науки. Поэтому наряду с термином «ступень наблюдения» возможно применение таких терминов, как «протокольная база науки» или «эмпирическая база науки».

Принципу двухступенчатости логики науки соответствует в кибернетике теория черного ящика. Согласно этой теории, объект познания есть своеобразный черный ящик, содержимое которого недоступно прямому наблюдению. Исследователь может составить гипотетическое представление о содержимом черного ящика, только подвергая его различным внешним воздействиям (входное состояние) и наблюдая за поведением черного ящика в результате этих воздействий (выходное состояние).

Мы полагаем, что последовательное проведение принципа двухступенчатости в структурной лингвистике освободит ее от целого ряда фундаментальных трудностей, с которыми она сталкивается в настоящее время. Речь идет не об априорном наложении схемы логики науки на структурную лингвистику. Напротив, внутри самой структурной лингвистики имеют место глубокие антиномии, необходимость преодоления которых неизбежно ведет к идее двух ступеней абстракции. Логика науки только помогает осмыслить эту необходимость.

В нашей книге «Проблемы теоретической фонологии» мы уже представили двухступенчатую систему понятий применительно к фонологии. В настоящее время мы считаем возможным представить двухступенчатую систему понятий для всей структурной лингвистики.

В самых общих чертах двухступенчатая система понятий структурной лингвистики представляется нам в следующем виде.

В фонологии мы расщепляем понятие дифференциального элемента и фонемы на две пары соотносительных понятий: на понятие дифферентора и дифферентоида и понятие фонемы и фонемоида.

Исходный пункт для построения двухступенчатой системы понятий в фонологии — необходимость преодолеть фундаментальные антиномии, возникающие в существующих теориях фонологии при исследовании фонологической системы языка. Остановимся на одной из этих антиномий. Мы имеем в виду антиномию транспозиции. Суть антиномии транспозиции заключается в следующем.

К числу основных положений, на которые опирается современная фонология, принадлежат:

1) фонемы есть элементы, служащие для различения лингвистических единиц;

2) фонемы есть звуковые элементы.

Но если принять эти положения, то из первого следует, что мы можем транспонировать фонемы из акустической субстанции в другие виды субстанции (в графическую, цветовую, тактильную и т. д.), а из второго следует прямо обратное: что мы не можем транспонировать фонемы из акустической в другие виды субстанции. Для преодоления данной антиномии мы должны расщепить понятие фонемы на понятие фонемы в собственном смысле и понятие фонемоида как физической субстанции, служащей субстратом фонемы. На основании аналогичного рассуждения мы расщепляем понятие дифференциального элемента на понятие дифферентора и дифферентоида. Фонемы и дифференторы принадлежат к ступени конструктов, а фонемоиды и дифферентоиды — к ступени наблюдения.

Дифференторы и фонемы — это абстрактные диакритические элементы, способные воплощаться в разных физических субстанциях: акустической, графической, цветовой, тактильной и т. д. Дифферентоиды и фонемоиды — это акустические элементы, воплощающие те или иные дифференторы и фонемы. Предлагается строго различать и не смешивать друг с другом: 1) акустические свойства и комплексы акустических свойств (звуки) как чисто физические понятия, 2) дифферентоиды и комплексы дифферентоидов (фонемоиды) как реляционно-физические понятия, 3) дифференторы и комплексы дифференторов (фонемы) как чисто реляционные понятия.

В грамматике мы расщепляем ее основные понятия — морфему и синтагму — на понятие морфемы и морфемоида, синтагмы и синтагмоида. С морфемоидами и синтагмоидами мы имеем дело на уровне субституции как операционного теста на ступени наблюдения. Что же касается морфем и синтагм, то они рассматриваются как чистые конструкты, которые постулируются нами для объяснения функционирования морфемоидов и синтагмоидов.

Необходимость расщепления понятия морфемы на понятие морфемы в собственном смысле и понятие морфемоида и понятия синтагмы на понятие синтагмы в собственном смысле и понятие синтагмоида обусловлена фундаментальными трудностями, на которые наталкиваются эти понятия при исследовании грамматической системы языка. Остановимся на некоторых из этих трудностей.

Одним из фундаментальных фактов всякого языка следует признать факт линейной упорядоченности лингвистических единиц. На это обратил внимание еще Ф. де Соссюр, который писал: «Означающее, будучи свойства слухового (аудитивного), развертывается только во времени и характеризуется заимствованными у времени признаками: а) оно представляет протяженность и б) эта протяженность лежит в одном измерении: это линия. Об этом совершенно очевидном принципе сплошь и рядом не упоминают вовсе,

по-видимому именно потому, что считают его чересчур простым; между тем это принцип основной и последствия его неисчислимы. От него зависит весь механизм языка» [1]. Линейность, являясь одним из фундаментальных свойств языка, обусловленных физическими особенностями языка, вступает в противоречие с внутренними синтагматическими связями лингвистических единиц. Это противоречие особенно ярко проявляется в существовании так называемых разрывных морфем.

Возьмем, например, английское предложение *He put it up*. Линейные свойства этого предложения вступают в противоречие с внутренней связью *put* и *up* как элементов единого целого. Так же обстоит дело и в конструкциях со вспомогательными глаголами, например во фразе *I have taken* линейные свойства данной фразы вступают в противоречие с внутренней связью *have* и *-en*, которые, будучи отделены друг от друга элементом *take*, представляют собой элементы единого целого.

Как же возможно преодолеть противоречие между линейными свойствами грамматических конструкций и внутренними связями элементов этих конструкций? Мы полагаем, что необходимо расщепить понятие морфемы на понятие морфемы как конструкта и понятие морфемоида как элемента, служащего субстратом морфемы. Так, в рассмотренных примерах единая морфема (*PUTUP*) воплощается в двух морфемоидах *put* и *up*, единая морфема (*HAVEEN*) воплощается в двух морфемоидах *have* и *-en*. Расщепление понятия морфемы на понятие морфемы в собственном смысле и понятие морфемоида влечет за собой, естественно, и расщепление понятия синтагмы на понятие синтагмы в собственном смысле и понятие синтагмоида. Так, в первом примере синтагма (*PUTUP IT*) воплощается в синтагмоиде *put it up*, а во втором примере синтагма (*HAVEEN TAKE*) воплощается в синтагмоиде *have taken*.

В связи с расщеплением понятия морфемы на понятие собственно морфемы и понятие морфемоида и понятия синтагмы на понятие собственно синтагмы и понятие синтагмоида возникает необходимость и в расщеплении понятия дистрибуции на понятие метрической дистрибуции и понятие линейной дистрибуции. Под линейной дистрибуцией мы имеем в виду дистрибуцию морфемоидов и синтагмоидов, тогда как под метрической дистрибуцией мы имеем в виду внутренние связи между морфемами и между синтагмами, независимые от каких-либо пространственных представлений.

Не следует смешивать понятие морфемоида с понятием морфа. В лингвистической литературе морфами принято называть кратчайшие грамматические элементы языка. Если различные по своему фонемному составу морфы находятся в отношении дополнительной дистрибуции друг к другу, то они объединяются в классы, называемые морфемами, и в качестве вариантов морфемы получают название алломорфов. Таким образом, морфемы и морфы относятся друг к другу как инварианты к своим вариантам. Но в нашем понимании морфем и морфемоидов они не относятся друг к другу как инварианты и варианты. Мы рассматриваем морфему и морфемоид как гетерогенные элементы, принадлежащие к разным ступеням абстракции, и поэтому проблема инвариантов и вариантов должна ставиться для каждой ступени абстракции раздельно. Рассмотрим конкретный пример.

Возьмем в русском языке фразы *уютная комната*, *комната уютна* и *сделать комнату уютной*. В этих фразах мы находим три конкретные морфемы прилагательных — атрибутивную морфему в слове *уютная*, которую обозначим символом A_1, предикативную морфему в слове *уютна*, которую обозначим символом A_2, и приобъектную морфему в слове *уютной*, которую обозначим символом A_3. Морфемы A_1, A_2, A_3 объединяются в класс адъективных морфем. Можно доказать, что основным параметром адъективных морфем является атрибутивный параметр, поэтому A_1 должна

[1] Ф. де Соссюр. Курс общей лингвистики, ч. 1, гл. 1, § 3. М., 1933.

считаться морфомой-инвариантом, а A_2 и A_3 должны считаться вариантами морфемы A_1.

Морфема A_1 воплощается в морфемоиде *-ая*, морфема A_2 — в морфемоиде *-а*, а морфема A_3 — в морфемоиде *-ой*. Морфемоиды *-ая*, *-а* и *-ой* объединяются в класс адъективных морфемоидов. Поскольку морфемоиды *-ая* и *-ой* могут трактоваться в качестве последовательных преобразований морфемоида *-а*, то этот последний должен считаться морфемоидом-инвариантом, а первые — его вариантами.

Таким образом, мы видим, что проблема инвариантов и вариантов должна решаться независимо для морфем и морфемоидов. Более того, между морфемами-инвариантами и морфемами-вариантами, с одной стороны, и морфемоидами-инвариантами и морфемоидами-вариантами, с другой стороны, отсутствует взаимно-однозначное соответствие: так, в нашем примере морфемоид-инвариант *-а* служит субстратом морфемы-варианта A_2, напротив морфемоид-вариант *-ая* служит субстратом морфемы-инварианта A_1.

Аналогичным образом проблема инвариантов и вариантов должна ставиться раздельно для фонем и фонемоидов. Возьмем пример из древнегреческого языка. В древнегреческом можно различать две конкретные фонемы: «s_1» (в позиции перед согласными — в группах типа *spr*, *sp*, *sm*) и «s_2» (в позиции перед гласными после смычных — в группах *ps*, *ks*). Фонема «s_1» воплощалась в фонемоиде $s>$, а фонема «s_2» — в фонемоиде $s<$ (диакритические значки $>$ и $<$ служат символами имплозивного и эксплозивного произношения). Если на фонемном уровне, опираясь на функциональные критерии, мы должны считать «s_1» фонемой-инвариантом, а «s_2» — вариантом фонемы «s_1», то на фонемоидном уровне мы должны считать инвариантом не фонемоид $s>$, а фонемоид $s<$, потому что с физической точки зрения наиболее назависимой является позиция эксплозивного произношения. Таким образом, мы видим на данном примере, что в фонологии отсутствует взаимно-однозначное соответствие между фонемами-инвариантами и фонемами-вариантами, с одной стороны, и фонемоидами-инвариантами и фонемоидами-вариантами, с другой. В самом деле, фонемоид-инвариант $s<$ служит субстратом фонемы-варианта «s_2», напротив, фонемоид-вариант $s>$ служит субстратом фонемы-инварианта «s_1».

В связи с различением двух ступеней абстракции мы считаем целесообразным ввести в структурную лингвистику два общих понятия — понятие лингвистического генотипа и понятие лингвистического фенотипа. Первое относится к ступени конструктов, а второе — к ступени наблюдения. Генотипы — это дифференторы, фонемы, морфемы, синтагмы. Фенотипы — это дифферентоиды, фонемоиды, морфемоиды, синтагмоиды и т. д. Мы считаем целесообразным назвать ступень конструктов генотипической ступенью, а ступень наблюдения — фенотипической ступенью.

Соотношение инвариантов и вариантов генотипической и фенотипической ступени можно представить в табл. 1.

Таблица 1

Планы языка / Ступени абстракции	Фонология		Грамматика	
Генотипическая	Фонема-инвариант	Фонема-вариант	Морфема-инвариант	Морфема-вариант
Фенотипическая	Фонемоид-инвариант	Фонемоид-вариант	Морфемоид-инвариант	Морфемоид-вариант

Такова в самых общих чертах двухступенчатая система лингвистических понятий.

Различение генотипической и фенотипической ступени абстракции проливает новый свет на теорию лингвистического моделирования. Систематическое рассмотрение следствий из этого различения для теории лингвистического моделирования мы пока отложим. Здесь же только укажем вскользь на один из аспектов проблемы моделирования, который представляется нам особенно существенным. Мы имеем в виду оценку модели непосредственно составляющих с точки зрения указанного различия двух ступеней абстракции.

В своих выдающихся работах Н. Хомский дал блестящую критику модели непосредственно составляющих, убедительно раскрыв сильные и слабые стороны ее. Как показал Н. Хомский, модель непосредственно составляющих должна быть дополнена трансформационной моделью. Как на один из основных недостатков модели непосредственно составляющих, Н. Хомский указывает на запрещение переставлять грамматические элементы, поскольку в силу этого запрещения модель непосредственно составляющих оказывается неприменимой к цепочкам с разрывными морфемами, которые во многих языках занимают центральное место в системе (например, в английском языке, где развита богатая система спряжения со вспомогательными глаголами, являющимися частью разрывных морфем). В качестве средства для преодоления этого недостатка Н. Хомский предлагает трансформации перестановок.

В свете принципа двухступенчатости можно поставить следующий вопрос: к какой ступени абстракции относится модель непосредственно составляющих?

И вот оказывается, что модель непосредственно составляющих привязана к фенотипической ступени. В самом деле, если в модели непосредственно составляющих запрещаются перестановки элементов, то, значит, для этой модели существенны пространственные представления, характерные именно для фенотипической ступени.

Дополняя модель непосредственно составляющих трансформациями перестановок, мы не поднимаемся над фенотипической ступенью. Поэтому если мы хотим построить гомогенную модель, имеющую дело прежде всего с генотипической ступенью, то для этой цели естественно искать модель нового типа, нежели модель непосредственно составляющих.

В качестве модели для исследования лингвистических генотипов мы предлагаем порождающую модель установления доминативных связей между элементами предложения путем так называемой операции аппликации. Описание аппликативной порождающей модели дано нами в другой работе [2]. Здесь только укажем, что существенной особенностью данной модели является полная свобода от пространственных представлений и учет лишь имманентных связей между лингвистическими элементами.

Поскольку аппликативная порождающая модель полностью свободна от пространственных представлений, то не возникает необходимости вводить в эту модель трансформации перестановок и другие трансформации, связанные с пространственными представлениями. Зато в этой модели играют весьма существенную роль трансформации, позволяющие установить отношение инвариантности между разными типами фраз и раскрыть глубинные связи между ними.

Таковы некоторые следствия из принципа двухступенчатости, касающиеся моделирования языка.

---

[2] С. К. Шаумян. Порождающая модель на базе принципа двухступенчатости. ВЯ, 1963, № 3. См. также С. К. Шаумян и П. А. Соболева. Аппликативная порождающая модель и исчисление трансформаций в русском языке. М., 1963.

Summary

The two-level principle which is the back-bone of the modern logic of science is suggested here as the logical basis for linguistic theory.

The problem of invarience arises separately at the two main levels of abstraction — the phenotype level and the genotype level.

In phonology we must distinguish between phonemes-invariants and phonemes-variants at the genotype level of abstraction, whereas at the phenotype level we must distinguish between phonemoids-invariants and phonemoids-variants.

In grammar morphemes-invariants and morphemes-variants must be distinguished at the genotype level, whereas at the phenotype level morphemoids-invariants and morphemoids-variants must be distinguished.

In connection with distinguishing between the phenotype and the genotype levels a necessity arises for a new generative linguistic model which the author calls an applicational generative model.

А. А. Зиновьев

ОБ ОСНОВАХ АБСТРАКТНОЙ ТЕОРИИ ЗНАКОВ

1. Среди множества различных направлений исследования знаков существенную роль играет направление, исходящее из интересов логики и из исследований философско-лингвистического толка, тесно связанных с идеями и течениями современной теоретической лингвистики. Мы имеем в виду работы Г. Фреге, Б. Рассела, У. Куайна, К. Люиса, А. Черча, Р. Карнапа, К. Айдукевича, Ч. Пирса, Ч. Морриса, Ф. де Соссюра и других авторов (см., например, [5]). При этом мы вовсе не склонны считать, что это есть некое единое направление, исходящее из общих всем его представителям принципов. Наоборот, даже в рамках работ логического толка легко заметить весьма существенные различия (см. об этом, например, в книге [4]). Мы объединяем эти весьма разнородные работы лишь постольку, поскольку им свойственна тенденция решать проблемы теории знаков в специфической для нее терминологии, решать их объективно (исключая все, что не поддается наблюдению и точному фиксированию), используя такие средства научного исследования, как идеализация, дедукция, формализация и т. п. Абстрактная теория знаков, о которой здесь пойдет речь, мыслится как теоретическая конструкция такого рода. Но в этой статье мы рассмотрим лишь некоторые общие условия ее построения, ее основания.

2. Под абстрактной теорией знаков мы имеем в виду теорию или, скорее, совокупность теорий, в которой исследуются гипотетические (абстрактно мыслимые, допускаемые) свойства, отношения и соединения любых знаков, не зависящие от физической природы последних и от характера их практического использования. Языковые знаки при этом выступают как частный случай знаков, хотя именно язык дает основной материал для иллюстраций, служит источником проблем и идей теории, постоянно имеется в виду как ближайшая (если вообще не единственная) сфера ее приложения. Если бы даже кроме языковых знаков не существовали никакие другие, такой абстрактный подход был бы все равно правомерен как средство отвлечения от множества обстоятельств и предварительного упрощения проблемы.

Проиллюстрируем сказанное. В абстрактной теории знаков возможно поступать, например, таким образом. Будем рассматривать соединения простых знаков в сложные. При этом нам о знаках известно только то, что они суть знаки. Типы соединений мы просто постулируем, например, таким образом: пусть $З^1$ и $З^2$ суть какие-то различные знаки; $З^1$ находится в соответствии $С^1$ с предметами множества $П^1$, а $З^2$ — в соответствии $С^2$ с $П^2$; $З^1 \cdot З^2$ пусть будет обозначать такое соединение $З^1$ и $З^2$ в новый знак, при котором множество предметов $П^3$, находящихся в соответствии $С^3$ с $З^1 \cdot З^2$, есть подмножество множества $П^1$ и подмножество множества $П^2$;

C^1, C^2 и C^3 могут быть тождественны или различны в различных комбинациях (см. об этом [3]). Конечно, вводя такое соединение знаков, мы предполагаем возможность его интерпретации, но в самой теории (формально) это введение выглядит как произвольное.

Из определения (и иллюстрирующего его примера) видно, что в абстрактной теории знаков не предполагается существования субъектов (предметных знаков, имен предметов и т. п.), предикатов, образованных из них высказываний и высказываний, образованных из высказываний. Не предполагаются и такие логические знаки, как «все», «некоторые», «возможно», «необходимо», «должно», «и», «или», «не» и т. п. Мы можем пользоваться этими знаками, поскольку не можем построить никакую теорию, не прибегая к высказываниям и рассуждениям. Но мы не рассматриваем всего этого в качестве исследуемого материала. Речь идет о любых знаках, и такие формы знаков, о которых только что говорилось, суть лишь частный случай. Более того, в абстрактной теории знаков лишь имеется возможность развить разделы, в которых будут введены понятия, допускающие интерпретацию как субъекты, предикаты, высказывания, соединения высказываний. Но это не будет непосредственное описание этих фактов — абстрактная теория знаков не заменяет теорию высказываний и теорию вывода.

Очевидно, что намеченный здесь путь прямо противоположен тому, который принят, например, в [4]: там совершается переход от субъектов, предикатов и высказываний к некоторым обобщениям, здесь же — от рассуждений о знаках вообще к таким их допускаемым формам, соединениям и применениям, которые можно интерпретировать как субъекты, предикаты и высказывания. Это различие в направлении хода мысли должно привести, надо думать, к существенным различиям в понимании проблематики теории знаков вообще.

Мы не случайно в начале этого раздела не упомянули об изменении знаков. Дело в том, что само выражение «изменение знака» нам представляется нелепым с точки зрения абстрактной теории знаков. Возьмем, например, слово «стол». Оно может измениться как физическое явление (выцветают чернила, разрушается краска; написано от руки и напечатано на машинке; напечатано курсивом и другим шрифтом и т. п.) и грамматически (склонение). Но оно неизменно в качестве знака столов, т. е. неизменен тот факт, что оно обозначает столы. Если оно начинает обозначать другие предметы и перестает обозначать столы, оно прекращает свое существование в качестве знака столов (или в качестве знака только столов, если просто расширяется его объем), но не меняется в самом этом качестве подобно тому, как растут деревья, меняется климат, меняется пространственная форма тела и т. п. Знаки вообще неизменны в их качестве знаков (если отвлечься от их физической природы и выделить в них только то, что они соответствуют каким-то предметам). Они могут появляться и исчезать, заменяться другими и т. п. Но если они существуют как знаки, они как таковые неизменны. Все это — неизбежное следствие самого подхода, именуемого здесь абстрактной теорией знаков. То, что порой называют изменением знаков, есть на самом деле изменение того материала, из которого построен знак, а это уже означает выход за пределы принятой абстракции.

Из сказанного следует чрезвычайно важный вывод, усиливающий правильное, на наш взгляд, утверждение Ельмслева в [2] о том, что языки не могут быть описаны как чисто знаковые системы: лингвистика, какой бы абстрактный характер она ни имела, вообще не есть теория знаков; она исследует строительный материал для знаков определенного рода, конструирование знаков (и не только знаков) из этого материала, изменения этого материала; лишь отчасти ее точка зрения на этот материал есть точка зрения теории знаков.

Термин «знак» употребляется в таких двух (среди прочих других) смыслах: в одном случае имеется в виду то, что предметы определенного рода отбираются или специально создаются для того, чтобы соответствовать другим, и рассматривается эта их специфическая роль; для лингвистики это не есть предпосылка ее теоретических конструкций; в другом случае под знаком имеется в виду все, что определенным образом пишется, произносится, делается (жест); в этом случае знаками считают такие явления, как «с», «а», «х», «у стола», «к столу», «—», «=» и т. п., не являющиеся знаками в первом смысле. Теорией знаков в первом смысле может быть лишь абстрактная теория знаков, в которой предметом исследования становится само свойство некоторых предметов (безразлично каких) быть знаками.

Такая теория не существует в строгом смысле слова «теория». Существует довольно значительная совокупность относящихся к знакам понятий, утверждений, предположений, рассуждений и т. п., которые можно найти в многочисленных лингвистических, логических, философских, исторических и т. п. работах, а также в исследованиях, синтезирующих точки зрения различных наук. В систематизации этого идейного материала некоторым строго формальным образом можно видеть одну из ближайших задач такой теории или, скорее, совокупности таких теорий.

В подавляющем большинстве случаев обсуждение проблемы теории знаков затрагивает лишь ее идейные основы (ее основные понятия, положения, методы) и ведется на описательно-содержательном, если можно так выразиться, уровне. Исследование на этом уровне необходимо как условие для отыскания наиболее перспективного (с той или иной точки зрения) формального подхода. Кроме того, оно имеет самостоятельный интерес для разрешения ряда дискуссионных проблем гносеологии, связанных со знаками. На этом уровне будут вестись и дальнейшие рассуждения.

3. Мы предполагаем, что читатель в какой-то мере уже знаком с рассматриваемыми здесь понятиями, имеет в отношении их интуитивную ясность и силу привычки оперировать знаками. Это относится прежде всего к самому термину «знак».

Однако это предположение выражает нечто большее, нежели только условия и цели данной статьи. Не говоря уже о том, что надо начинать с чего-то уже известного, в рамках абстрактной теории знаков понятие «знак» вообще может быть принято как первичное, без эксплицитного определения типа: «Знаком будем называть . . .». Мы не исключаем возможности экспликации этого понятия. Но здесь в этом нет необходимости: в самой теории знаков такое понятие о знаке не играет никакой формальной роли; для содержательного подхода достаточно интуитивного понимания или разъяснений такого рода, какие мы дадим ниже; этих разъяснений будет вполне достаточно, чтобы выяснить ориентацию теории знаков и необходимые для ее развития исходные понятия. Знаки — объект исследования теории знаков. И этот объект дан нам эмпирически в виде множества случаев употребления, конструирования и т. п. знаков подобно тому, как ботанику даны растения, экономисту — товары, физику — атомы и т. п. Определение этого объекта по схеме «Знак есть. . .» («Знаком называется . . .») будет играть скорее роль предварительного описания или краткого резюме исследования, а не формально необходимого элемента теории.

Но обратимся к разъяснению термина «знак». Для этого надо сначала разъяснить другие термины, а именно — термины «предмет» и «соответствие». С первым из них дело обстоит просто: предметами мы называем не только то, чему соответствуют предметные имена или субъекты высказываний (суждений), но и все то, чему соответствуют также предикаты и

целые высказывания. Этой ссылки на субъекты, предикаты и высказывания лучше не делать, так как того же эффекта можно добиться перечислением примеров предметов, обобщение которых вполне достаточно для крайне общего смысла термина «предмет»: предмет — стол, атом, заряд, желтое, цвет, число, пять, стоимость; тот факт, что пять больше трех; тот факт, что Ленинград севернее Москвы, а Москва севернее Киева; тот факт, что масса и скорость тела связаны определенным соотношением; флогистон, слово, кентавр и т. п. Такого рода перечисление конечного числа разнородных примеров вполне гарантирует выработку такого навыка оперирования словом «предмет», когда этим словом можно назвать абсолютно все, что угодно. Заметим, между прочим, что этот прием введения общих понятий широко распространен в науке, но не зафиксирован в теории определений.

Сложнее обстоит дело с соответствием (см. об этом [3]). Здесь приходится предполагать известным, что́ такое ощущение, восприятие и т. п., т. е. предполагать данной целую совокупность понятий. Но без этого не обойтись, и ощущения, восприятия и т. п. придется рассматривать как факты, известные всякому среднекультурному человеку из его опыта. Для сокращения можно ввести термин «выбор»: выбрать предмет — ощутить, воспринять, вспомнить (представить), вообразить, назвать его, а также высказать о нем некоторую совокупность высказываний, каким-то образом изобразить его и т. д. Здесь нет угрозы порочного круга, так как мы всю эту терминологию берем как привычно ясную и используем ее для описания столь же привычных фактов, а не для формально строгих конструкций. Установить соответствие предмета $П^1$ предмету $П^2$ — значит добиться следующего (будем так считать): при некоторых в принципе точно учитываемых условиях вслед за выбором $П^1$ осуществляется выбор $П^2$. Или другими словами: имеются такие условия, что всегда вслед за выбором $П^1$ выбирается $П^2$. Например, если мы условились, что выражению «x истинно» соответствует «контакт a замкнут», то при необходимости интерпретировать некоторую формулу в терминах второго выражения мы вслед за выбором термина «истинно» выбираем термин «замкнут».

Условия, в которых шла речь, в каждом случае должны быть заданы и не могут быть обобщены в некоторой стандартной формулировке, за исключением ссылки на конвенцию и на необходимость воспользоваться ею, о чем скажем ниже. Напоминаем, что мы все время движемся на описательном уровне, так что данное определение можно расценивать как указание на некоторые черты соответствия, важные в данном контексте. Вместо выражения со словами «вслед за» можно сказать «если выбирается $П^1$, то (при некоторых условиях) выбирается $П^2$». Но в этом нет нужды: условия можно понимать так, что «вслед» будет иметь силу «вследствие». Возможно соответствие $П^2$ предмету $П^1$ (т. е. в другую сторону) и соответствие в обе стороны. Если $П^1$ и $П^2$ берутся как представители множеств, то достаточно говорить: «Любой $П^1$ соответствует любому $П^2$». Говоря для простоты о соответствии одного предмета другому, мы будем иметь в виду соответствие любого предмета одного множества любому предмету другого. Кроме того, будем говорить о соответствии $П^1$ и $П^2$, отвлекаясь от того, соответствует $П^1$ предмету $П^2$, $П^2$ предмету $П^1$ или и то и другое. Будем записывать это символами типа $П^1$—$П^2$. Ряд свойств соответствия укажем ниже.

Теперь можно дать такое определение знака: если специальной функцией предмета $П^1$ является то, что он вступает в соответствие (находится в соответствии) с предметом $П^2$, то $П^1$ называется знаком для $П^2$, а $П^2$ — обозначаемым для $П^1$. В дальнейшем знаки будем обозначать символами $З$, $З^1$, $З^2$, . . ., а символы $П$, $П^1$, $П^2$, . . . сохраним за обозначаемыми. Какие предметы начинают играть роль знаков и почему, этот вопрос

лежит вне компетенции абстрактной теории знаков. Термин «соответствие» можно не разъяснять, как было сделано выше, и принять за первичный. Но это не избавляет от необходимости разъяснений какого-то иного рода.

Данное определение является точно так же описательным. Однако оно играет ту существенно важную роль, что дает ориентацию для дальнейших рассуждений: дает содержательное описание исходного пункта этих рассуждений, а именно — описание схемы $З—П$.

В качестве исходного пункта схема $З—П$ связана с целым рядом абстракций, которые частично учтены в самих данных выше определениях и частично могут быть выявлены на их основе. Рассмотрим их подробнее. Одновременно будем отмечать некоторые свойства рассматриваемых объектов, представляющие интерес.

4. Соответствие предметов устанавливается не само по себе, а человеком. Без человека (вообще без существа, способного осуществлять выбор предметов) ни о каком соответствии и речи быть не может. Так что в схему $П^1—П^2$ должен быть добавлен третий элемент — человек, и схема примет вид, допустим, такой: $П^1 \rightarrow Ч \leftarrow П^2$, где стрелки указывают только на факт выбора. Эта схема точнее первой, поскольку первая создает видимость связи $П^1$ и $П^2$ независимо от человека.

Конечно, связь предметов $П^1$ и $П^2$ независимо от человека вполне возможна. Но когда речь идет о соответствии, то определяющим является следующее. Процесс выбора предмета не влияет на него сколько-нибудь заметным образом, так что практически можно сказать, что выбор оставляет предмет неизменным (например, назвав $П$ каким-то именем, мы тем самым не меняем его). Это касается обоих предметов, между которыми устанавливается соответствие. Более того, при этом также исключается влияние их друг на друга. Если это влияние (взаимодействие) есть, то оно с точки зрения соответствия $П^1$ и $П^2$ не существенно, не меняет предметы в том их качестве, в каком они суть $П^1$ и $П^2$. В противном случае это будут другие предметы, это не будет сопоставление $П^1$ и $П^2$. Именно потому, что человек осуществляет выбор $П^1$ и $П^2$, он фиксирует в них нечто неизменное и тем самым исключает их взаимодействие. Практически это выглядит так, что имеет смысл устанавливать соответствие $П^1$ и $П^2$ тогда, когда они на самом деле сравнительно устойчивы и практически не взаимодействуют (изменениями и взаимодействием можно пренебречь). Сказанного достаточно, чтобы сделать вывод об отличии $П^1—П^2$ от причинной связи: последняя берется как нечто независимое от человека, выделяется именно изменение предметов в зависимости друг от друга (в частности — их возникновение или уничтожение).

Определение соответствия, которое мы дали, указывает не на то, как создаются условия для $П^1—П^2$, а на сам акт осуществления $П^1—П^2$ при данных уже условиях, на сам познавательный процесс перехода от одного предмета к другому. Этот акт может повторяться (без расчета на это соответствие теряет смысл) в стандартных условиях, последние же устанавливаются как нечто постоянно действующее (в определенных временных рамках, конечно). Среди них необходимо указать на наличие конвенции о соответствии $П^1$ и $П^2$. Эта конвенция может иметь стихийно-бессознательный характер, может формулироваться также вполне сознательно в виде соглашений «Пусть $П^1—П^2$». Обращаем внимание на то, что положения «Пусть $П^1—П^2$» (а также стихийные соглашения, в принципе допускающие аналогичное фиксирование) не образуют самого соответствия в рассматриваемом нами смысле, а лишь образуют одно из его условий. Возьмем такой пример. В некотором процессе познания мы приняли решение: пусть объекту x одного множества объектов соответствует объект y другого множества объектов. Это соглашение еще не есть соответствие как некоторый акт. В ходе исследования объектов первого множества мы получили какой-то результат, который надо исполь-

зовать для понимания объектов второго множества. Тут-то мы и вспоминаем о нашей конвенции и осуществляем, используя ее (и еще какие-то дополнительные данные), переход к объектам второго множества, в том числе от x к y. О том, что соответствие установлено, мы и судим по этому переходу. Конечно, положение «Пусть $П^1$—$П^2$» достаточно ясно говорит о наличии $П^1$—$П^2$. Но если потребуется расшифровать кому-то термин «соответствие», то придется пойти тем путем, который мы привели здесь. А чтобы выделить соответствие в чистом виде, придется постулировать условия, в которых последовательность выбора равносильна следованию. Впрочем, указание на наличие конвенции «Пусть $П^1$—$П^2$» и на необходимость воспользоваться ею в каком-то познавательном процессе достаточно для того, чтобы это стало возможно. То, что мы так много говорим о соответствии (а где же знаки?!), пусть не смущает читателя: достаточно один из $П^1$ и $П^2$ принять в качестве знака, как все сказанное будет сказано о знаках.

Обратимся к схемам $З$—$П$ и $З \rightarrow Ч \leftarrow П$. Физическая природа знаков в них (и соответствующих определениях) задана тем и только тем, что выбирается определенная символика для их обозначения — $З$, $З^1$, $З^2$, а также символика для сложных знаков. Имеется в виду любой знак и любое обозначаемое, т. е. предполагается рассматривать само отношение $З$—$П$ и только. Предполагая любой $П$ (не в смысле «каждый», а в смысле «какой-то», «безразлично, какой»), мы тем самым отвлекаемся от проблемы существования $П$: существует $П$ или нет, роли не играет; достаточно того, что мы всегда можем без каких бы то ни было затруднений высказать «Пусть дан $П$», «$З$ соответствует $П$» и т. п. Что касается $З$, то он всегда в принципе может быть дан в ощущениях и восприятиях (непосредственно чувственным образом), и проблема существования для него вообще лишена смысла. Это положение вполне согласуется с эмпирическими фактами: люди специально создают знаки такими, чтобы их можно было ощутить и воспринять, так как без этого невозможно пользоваться ими и передавать (сообщать) их другим лицам. Это обстоятельство позволяет в дальнейшем не обращать внимания на то, что при установлении $З$—$П$ знак дан человеку в виде чувственного образа, т. е. отвлечься от духовного (как иногда говорят) плана проблемы и говорить о знаке как об объективном факте.

Схема $З$—$П$ говорит лишь о том, что как-то установлено соответствие $З$ и $П$ (что знак отнесен к предмету, соотнесен с предметом и т. п.). Для некоторых целей может быть этого достаточно. Если это так, то содержательные разъяснения можно свести к описанию навыка употребления знаков, например, следующим образом: соответствие $З$ и $П$ будем считать установленным, если данный человек может указать предмет, обозначаемый данным знаком, или указать знак, обозначающий данный предмет, или сделать и то и другое; здесь слово «знак» всегда связывается с явлениями определенной физической природы и речь идет о знаке $З^1$, знаке $З^2$ и т. п., а не о знаке вообще. Но в теории знаков для более четкого определения ее предмета этого, надо думать, недостаточно. Необходимо рассмотреть, как устанавливается $З$—$П$, т. е. рассмотреть схему $З \rightarrow Ч \leftarrow П$ с учетом данного выше определения соответствия и перечисленных его свойств.

Прежде чем перейти к дальнейшей детализации, сделаем несколько разрозненных замечаний. Очевидно, что между $З$ и $П$ в $З$—$П$ нет причинной связи в том смысле, что один из них не есть причина появления, уничтожения или вообще изменения другого. Если бы $З$—$П$ было причинной связью, то $З$ не смог бы выполнить функцию знака: только начав выполнять эту функцию, он либо сам исчез бы и изменился (превратился в нечто иное), либо уничтожил бы или изменил $П$ так, что он перестал бы быть $П$; т. е. начало $З$—$П$ было бы одновременно и концом. Практически $З$

выбирается так, чтобы исключалось заметное воздействие *З* и *П* друг на друга.

Одно из следствий смешения причинности и соответствия состоит в том, что утверждение «Предметы различны (тождественны), если различны (тождественны) обозначающие их знаки», вполне правомерное как постулат какого-либо раздела или как следствие других положений (например, положения о взаимно-однозначном соответствии *З* и *П*) теории знаков, воспринимается как ошибочное утверждение о том, что тождество и различие знаков есть причина тождества и различия предметов. А между тем в нем речь идет либо об определении понятий тождества и различия предметов, либо о взаимно-однозначном соответствии *З* и *П* и т. д., но уж никак не о причинной зависимости. Аналогичные недоразумения возникают в отношении других утверждений.

На основе сказанного легко различить знак и чувственный образ. Последний есть состояние мозга и связанной с ним отражательной системы (аппарата отражения, скажем), неотделимый от них. Тогда как знак не есть состояние аппарата отражения, он является внешним по отношению к этому аппарату. Знак не есть чувственный образ. Он сам может отображаться чувственно. Чувственный образ не есть знак и никогда не может быть таковым по отношению к аппарату отражения, в котором он возникает. Он есть средство оперирования знаками, но не знак. Проблема же отношения знака и чувственного образа, на наш взгляд, не есть достаточно трудная проблема, если договориться о смысле соответствующих понятий.

Правильнее говорить не об отношении «знак — предмет», а об отношении «знак — обозначаемое», ибо знак есть сам предмет, который лишь в силу особой роли в некоторой системе связей приобретает функцию, благодаря которой мы и даем ему особое название «знак». Знак есть, повторяем, предмет, находящийся в особого рода отношениях. И никаких идеальных сторон в нем самом нет и быть не может, если только мы на самом деле стремимся к контролю за всеми элементами изучаемой действительности.

5. Будем называть чувственным образом *П* его ощущение, восприятие, представление, воображение; причем, безразлично, как возник чувственный образ *П* — в результате непосредственного воздействия или косвенно (например, через рисунок). Совокупность знаков, соответствующих *П*, будем называть знаковым отображением или описанием *П*. Это может быть простой знак и совокупность знаков, образующих сложный знак. Как чувственный образ *П*, так и его знаковое описание суть явления, поддающиеся наблюдению и строго объективному изучению. Соответствие *З* и *П* устанавливается либо посредством первого, либо посредством второго, либо посредством обоих. Схема *З—П* и выступает как сокращение того, что соответствие как-то установлено. С субъективной точки зрения это — ассоциация образов *З*, *П* и знакового описания *П*. Но эта точка зрения неприемлема, поскольку при этом исчезает специфика проблемы, а именно — свойства, отношения и связи знаков как объективных явлений. Чувственные образы — не предмет исследования абстрактной теории знаков. Мы лишь предполагаем их, вообще предполагаем, что конечной основой знакотворчества является множество чувственных образов предметов. Собственный объект абстрактной теории знаков — установление соответствия *З* и *П* посредством знакового отображения (описания) *П* и, на этой основе, правила построения сложных знаков, правила введения новых знаков с помощью имеющихся знаков и т. п.

Тот факт, что в определении знака скрытым образом подразумевается уже понятие знака, не означает порочного круга. Дело в том, что мы здесь имеем дело с определением, напоминающим определения индуктивного (рекурсивного) типа: 1) определяем частный случай соответствия с по-

мощью только понятия чувственного образа (о знаках и даже о словах и высказываниях не упоминаем); 2) предмет $П^1$, специально предназначенный для соответствия такого рода с $П^2$ и только для этого, называем знаком; 3) расширяем понятие соответствия за счет расширения понятия выбора (включаем сюда знаки, указанные в пункте 2); 4) в число знаков зачисляем предмет $П^1$, специально предназначенный для соответствия с $П^2$ (в расширенном смысле и только для этого); 5) полагаем, что никакие другие предметы знаками не являются. Это определение соответствует в какой-то мере истории познания и формированию индивидуального познания (с точки зрения формирования багажа знаков человека и человечества).

Перейдем к термину «значение знака». Здесь возможны различные точки зрения, в частности такая: значение данного знака состоит в том, что он соответствует чему-то отличному от него, обозначает какие-то предметы. Причем имеется способ выяснить, какие именно предметы (предмет) обозначаются. Это может быть указание на чувственно воспринимаемые предметы, описание последних совокупностью слов и предложений и т. п. Но во всех случаях поддающимся наблюдению и объективному изучению будет либо чувственный образ того предмета (состояние отражательного аппарата), с которым сопоставляется знак, либо его знаковое описание. И это обстоятельство дает право на иную точку зрения, о которой скажем ниже.

Знак имеет или не имеет значение лишь с точки зрения представителя какого-то множества людей, а не вообще. Знак имеет значение (можно так условиться), если представитель этого множества людей может каким-либо способом выделить из множества предметов по крайней мере один такой, который находится в соответствии с этим знаком, или дать описание такого предмета с помощью других знаков, имеющих значение. В частности, знак имеет значение, если по договоренности считается тождественным по значению другому знаку, имеющему значение. Знак не имеет значения, если все указанные выше операции не могут быть осуществлены (и тогда он вообще не знак). Мы допускаем, что в каждом частном случае некто (тот, для кого это важно) может каким-то способом установить значение данного знака или выяснить, имеет он значение или нет.

Если принять изложенную точку зрения, то изучение значений будет выступать как изучение способов установления значений, а именно — чувственных образов предметов (не входит в задачу теории знаков) и их знаковых описаний (входит в задачу теории знаков). Но к тем же результатам мы придем, приняв несколько иную систему понятий, а именно: если установлено $З—П$, то чувственный образ $П$ будем называть чувственным значением $З$, а знаковое описание $П$ — знаковым значением $З$. Схема $З—П$, таким образом, будет означать, что знак имеет либо первое, либо второе, либо оба значения. Если знак не имеет ни того, ни другого, то нет $З—П$ и $З$ не есть знак.

Значение знака при этом не есть нечто, обнаруживаемое в нем самом. Оно вне знака. Но это и не отношение знака к предмету или его отображению, это — элемент отношения «знак — чувственный образ или знаковое отображение предмета, находящегося в соответствии с этим знаком». Поскольку в абстрактной теории знаков чувственное значение предполагается, но не исследуется, то вполне допустимой с этой точки зрения оказывается абстракция, при которой в качестве значения признается только знаковое значение и отношение знака и значения расценивается как отношение знака к знаку. Если знак $З^1$ не имеет чувственного значения, то последнее есть соглашение считать другой знак $З^2$ значением первого (это — обобщение определения терминов). Семантическая проблематика здесь перерастает в своего рода синтаксическую,

оставляя семантический аспект лишь в том, что в теории охватывается символами типа *З—П*.

Значение знака не есть нечто идеальное в философском смысле: это — вполне поддающиеся объективному наблюдению состояния человеческого отражательного аппарата (объект изучения психологии и физиологии) и знаки (объект теории знаков). Как знаки они — не внутреннее состояние мозга и связанной с ним системы, а внешние по отношению к нему явления, чувственно отражаемые предметы. Очевидно, что в отличие от субъективного подхода, когда проблема переносится в план отношения духовных явлений, излагаемый здесь подход можно назвать объективным. Во втором случае духовный план предполагается как условие существования объективного плана (например, определение термина невозможно без отражательной деятельности человека, но если оно построено, то мы имеем вполне объективный комплекс знаков, который можно изучать как естественно данное явление на общих основаниях опытных наук).

Знаки можно сравнивать. Это можно делать прежде всего по их воспринимаемому виду, по их физической природе. На этот счет надо допустить у человека способность идентифицировать и различать знаки с этой точки зрения, т. е. принять выражения «знаки тождественны (различны) физически» или просто «знаки тождественны (различны)» как первичные и привычно ясные. Это относится и к сложным знакам, т. е. к знакам, конструируемым из других знаков: предполагается способность анализа сложных знаков и различения их с точки зрения воспринимаемой структуры (в структуре важно и то, какие знаки суть элементы сложного; принципы анализа должны быть указаны в форме правил построения сложных знаков). Но, как увидим ниже, понятий о физическом тождестве и различии знаков не всегда достаточно для сравнения знаков, и приходится говорить о позиционном тождестве и различии.

Тождество и различие знаков по чувственному значению определяется через тождество и различие их чувственных значений, а тождество и различие по знаковому значению — через тождество и различие знаковых значений. Здесь сложности никакой нет. В общем виде можно сказать: два знака тождественны по значению, если каждое значение одного из них есть также значение другого, и различны по значению, если среди значений одного из них найдется хотя бы одно значение, не являющееся значением другого. Если допускается только одно значение у знака, то получим частный случай этой формулировки: знаки тождественны (различны) по значению, если тождественны (различны) их значения.

Если знаки тождественны, то в соответствии с той абстракцией, о которой говорилось в разделе 3, их можно считать экземплярами, повторением, воспроизведением и т. п. одного и того же знака. И наоборот, различные экземпляры одного знака можно считать тождественными по значению знаками. Так что возможности для преодоления кажущихся трудностей гносеологического порядка имеются уже в самых простых определениях: можно начать с рассмотрения единичных случаев знаков, ввести понятие о тождестве знаков по значению и говорить о множестве тождественных по значению знаков как об одном абстрактном знаке; если не устраивает наличие нескольких знаков с одинаковым значением, их можно брать как экземпляры одного абстрактного знака **и т. п.**

Множественность значений знака имеет разнообразные познавательные причины (знак обозначает различные предметы, различные стороны одного и того же предмета, разные этапы познания предмета и т. п.). Если факт множественности значений важен, то в теории знаков это может быть учтено путем соответствующей конструкции определений (см., например, выше определение тождества и различия знаков по значению), в классификации значений и в выяснении их взаимоотношений (исклю-

БИБЛИОТЕКА
ИН-ТА РУССКОГО ЯЗЫКА
АН СССР

www.ruslang.ru

чают они друг друга или нет; если не исключают, то каковы последствия их соединения; можно ли выделить значения, каждого из которых достаточно для введения данного знака, и можно ли какое-либо из них признать за единственное значение, отбросив остальные, и т. п.).

Множественности значений в рамках теории можно «избежать», введя понятие позиционного различия знаков. Позиция знака — множество сложных знаков, в каждом из которых данный знак имеет только одно данное значение (напоминаем, что все определения такого рода мы вводим как иллюстрации, но не как нечто обязательное и единственно возможное). Если возможно несколько позиций данного знака, то можно говорить о позиционном различии физически тождественных знаков, ввести обобщенное понятие различия, охватывающее физическое и позиционное различие знаков, и оперировать уже с ним. Если знак имеет чувственное и знаковое значения, устанавливающие соответствие его с одним и тем же предметом, то первое из них можно понимать как интерпретацию второго; если ж они устанавливают соответствие знака с разными предметами, то их можно считать различными; впрочем, можно ограничиться рассмотрением только знаковых значений, полагая, что по части чувственных значений никаких осложнений не вносится.

Если позиции двух физически различных знаков тождественны во всем, кроме самих этих знаков, так что позиция одного может быть получена путем замещения им другого знака в позиции последнего, то эти знаки позиционно тождественны. Теперь можно ввести обобщенное понятие тождества знаков, охватывающее физическое и позиционное тождество, и оперировать с ним. Все это возможно сделать, надо думать, на строго формальном уровне. Если множества значений двух знаков пересекаются, то для каждого знака действует понятие позиционного различия, а для части позиционно различных знаков — понятие позиционного тождества; при полном совпадении множеств значений имеем позиционно различные пары позиционно тождественных знаков. Отметим, наконец, возможность введения вероятностей.

6. Мы до сих пор ничего не говорили об объеме знака. Понятие объема, конечно, полезно в теории знака. В частности, через отношение классов предметов, соответствующих данным знакам, просто характеризуются также отношения этих знаков как включение и взаимное включение (в применении к терминам — более или менее общий термин, термины одинаковой степени общности). Но с понятием объема связан целый ряд трудностей, поскольку приходится расширять круг онтологических понятий. Мы здесь подробно этот вопрос рассматривать не будем. Ограничимся краткими замечаниями.

Прежде всего следует отметить двоякий смысл термина «объем» знака: 1) в первом смысле имеется в виду реальное существование предметов, находящихся в соответствии с данным знаком, и множество этих предметов; в этом смысле можно говорить о пустых и непустых знаках; 2) во втором смысле имеется в виду лишь то, что с данным знаком связывается какое-то множество предметов, причем вопрос об их реальном существовании не ставится (например, мы считаем два знака «кентавр» и «человеко-лошадь» тождественными по объему независимо от того, существуют кентавры или нет); в этом смысле отношение объемов есть лишь удобное средство для описания отношения значений. В первом смысле проблема выходит за рамки теории знаков. Во втором ее можно элиминировать таким, например, способом (это, конечно, не единственный способ).

Среди значений *З* имеется такое, которое изображается схемой *З—П*. Это, в частности, может быть чувственное значение *З* или такое знаковое значение, знаки которого имеют уже чувственное значение; короче говоря,

это может быть некоторая интерпретация знака. И такое привилегированное (можно сказать, объемное) значение допускается для всех знаков. С таким значением можем связать понятие объема: если $З^1$—$П^1$ и $З^2$—$П^1$ и исключаются $З^1$—$П^2$ и $З^2$—$П^2$, т. е. для $З^1$ и $З^2$ исключаются другие объемные значения, то $З^1$ и $З^2$ тождественны по объему; если $З^1$—$П^1$ и $З^1$—$П^2$, а для $З^2$ имеет место только $З^2$—$П^1$ (или $З^2$—$П^2$), то $З^1$ шире по объему, чем $З^2$; возможны и другие отношения (см., например, [2]). На этом пути сохраняются все соотношения объема и значения в обычном смысле (например, тождество знаков по объему или по объемному значению не исключает различие по другим значениям). Но вообще-то говоря, различие значений знака, представляемое схемами $З^1$—$П^1$ и $З^1$—$З^2$ (последняя есть соглашение считать $З^2$ значением $З^1$), существенно важно: первая есть лишь некоторое сокращающее средство отождествления и различения знаков и анализу здесь не подлежит, вторая же должна быть специально исследована (построение знаков из знаков, введение знаков на базе других знаков).

Интересно было бы исследовать специально вопрос о возможности и путях определения ряда онтологических понятий через понятия теории знаков. Например, мыслимы такие (ориентировочно) определения: 1) если $З^1$ и $З^2$ различны (тождественны), то (в случае единственности значения знака) различны (тождественны) соответствующие предметы, т. е. соответствующие предметы называются различными (тождественными); 2) если $З^1$ и $З^2$ различны (тождественны) позиционно, то соответствующие им предметы называются различными (тождественными); 3) если $З^1$ шире по объему $З^2$, то множество предметов, соответствующих $З^2$, есть собственное подмножество множества предметов, соответствующих $З^1$, и т. п.

7. Если отвлечься от изменения материала для знаков при построении сложного знака, то последний можно определить как упорядоченную некоторым образом совокупность знаков, которая сама есть знак. С учетом возможности изменения материала для знаков определение можно несколько «ослабить»: сложный знак — знак, при построении которого использовались другие знаки. Но первое определение лучше в том смысле, что оно указывает на необходимость соглашения об упорядоченности знаков и о том, что упорядоченное таким образом множество знаков есть знак (первое соглашение выражается в структуре знака, второе имеет чисто прагматический характер). Это сразу дает ориентацию для формальных определений. Например, если уже определен перечень символов, то возможно такое определение знака: 1) $З$, $З^1$, $З^2$, . . . суть знаки; 2) если $З^1$ и $З^2$ знаки, то $З^1+З^2$, $З^1 \cdot З^2$ и $З^1 : З^2$ суть знаки; 3) никакие другие знаки не существуют. Здесь символы «+», « · » и « : » указывают на типы соединения знаков в сложные знаки. Для них принимаются утверждения, например, такого типа: если $З^1$—$П^1$ и $З^2$—$П^2$, то $З^1 \cdot З^2 > З^1$ и $З^1 \cdot З^2 > З^2$; $З^1 \cdot З^2 = З^2 \cdot З^1$ и т. п. Число типов соединений и характер утверждений для них в принципе не детерминированы, но, конечно, интересы интерпретаций здесь играют роль — влияют на абстрактные допущения. Заметим, что при формулировке правил соединений знаков в сложные никаких содержательных ограничений на значения знаков в самой теории знаков накладывать нельзя. При этом, например, такие соединения, как «железное дерево», «деревянное дерево» и т. п., могут оказаться правильно построенными знаками. Но это не должно смущать: общая теория рассматриваемого типа не есть абсолютно точное описание фактов, но лишь методическая рекомендация для их изучения.

Построение сложного знака есть введение нового знака особым образом, отличным от введения, которое можно представить как обобщение определения (см. об этом [3]). При этом требуется соглашение об отно-

шении значений (объемов в том числе, если имеется в виду объемное значение) вводимого знака и уже имеющегося знака. Например, это соглашение имеет такой вид: пусть $З^1$ находится в таком же соответствии с $П$, как $З^2$. Изложение такого рода соглашений в виде некоторой теории может быть интерпретировано, в частности, как теория определений терминов.

Требуется ввести корректив в понимание значения знака с учетом того, что он может быть сложным. Пусть $\{З^1, З^2\}$ есть сложный знак, образованный из $З^1$ и $З^2$, $З^{1*}$ и $З^{2*}$ — знаковые значения $З^1$ и $З^2$. Знаковым значением $\{З^1, З^2\}$ будут, можно так договориться, знаки $\{З^{1*}, З^2\}$, $\{З^1, З^{2*}\}$ и $\{З^{1*}, З^{2*}\}$ (здесь символы указывают лишь на то, что знак построен из фигурирующих в скобках знаков). Этот процесс можно продолжить и далее, пока не дойдем до чувственных значений (так что можно говорить о разветвлениях и степенях значений сложного знака). Все степени и разветвления получаемых при этом знаков тождественны по значению (это, очевидно, точно так же абстрактные допущения сравнительно с эмпирическими фактами). Практически же этот процесс останавливается в какой-то точке, которая считается привычно ясной или достаточной в данных условиях.

Два знака $\{З^1, З^2, \ldots\}$ и $\{З_1, З_2, \ldots\}$ будем называть тождественными по значению, если каждому составляющему знаку одного будет соответствовать составляющий знак другого, тождественный ему по значению; будем их называть различными по значению, если среди составляющих знаков одного окажется такой, что не найдется соответствующего составляющего знака другого, тождественного ему по значению. Тождество и различие здесь уже определены в общем виде, т. е. для элементарных знаков. Сам переход к определениям для сложных знаков можно формально выразить точно так же в виде индуктивных определений. Между прочим, не исключена возможность, что различные в некоторой степени или в некотором ответвлении по значению сложные знаки могут оказаться тождественными в другой степени или ответвлении, а тождественные — различными.

Развитие теории знаков можно осуществлять далее по линии исследования различных уровней знаков: знак — знак знака — знак знака знака и т. п. При этом на базе принятых определений очень простое обоснование получают утверждения об оперировании знаками разных уровней (или типов), например такие: знак никогда не может быть знаком самого себя, знак $З$ и знак этого знака («$З$») никогда не тождественны по значению и т. д. (даже тогда, когда мы имеем слово «знак» и обозначающий его знак «„знак“», на базе определений легко показать различие их значений). Не исключена возможность того, что многие семантические парадоксы могут получить здесь более простое разъяснение.

При рассмотрении знаков различных уровней можно говорить о таких правилах соединения знаков, которые можно интерпретировать как правила логики высказываний. Например, это будут знаки «Предмет, обозначаемый знаками $З^1$ и $З^2$», «Предмет, обозначаемый знаком $З^1$ или знаком $З^2$ (любым одним из знаков $З^1$ и $З^2$)» и т. п.

Мы вводили целый ряд понятий таким образом, что они оказались совершенно независимыми от различного рода утверждений о соотношениях знаков. Это удобно, поскольку оставляет свободу для различных направлений теории, базирующихся на различных (и подчас взаимоисключающих) утверждениях. Например, все определения тождества и различия даны независимо от принципов замены одних знаков другими; благодаря этому без изменения определений допустимы различные принципы замещения: пусть $\{З^1, \ldots\}$ и $\{З^2, \ldots\}$ образуются один из другого путем замещения одного из $З^1$ и $З^2$ другим; если $З^1$ и $З^2$ физически тождественны, то $\{З^1, \ldots\}$ и $\{З^2, \ldots\}$ тождественны по значению; если $З^1$

и $З^2$ позиционно тождественны, то $\{З^1, \ldots\}$ и $\{З^2, \ldots\}$ тождественны по значению при том условии, что $З^2$ не входит в $\{З^1, \ldots\}$, а $З^1$ не входит в $\{З^2, \ldots\}$; и т. п. Согласно одним из утверждений такого рода, оказывается допустимой, например, замена термина «столица СССР» термином «Москва» в предложении «Москва — столица СССР», и мы получим «парадоксальные» результаты (см. [6]); согласно другим, такая замена исключается (для чего достаточно потребовать, чтобы замещающий знак не входил в тот знак, в который входит замещаемый).

8. Выше мы договорились считать знак сложным, если могут быть обнаружены другие знаки, отличные от него по значению, из которых он построен. Очевидно, знак будет простым, если такие знаки не могут быть найдены (простой знак построен из самого себя). Сложный знак строится в конечном счете из двух или более простых по определенным правилам. Мы опять-таки допускаем здесь наличие навыка, позволяющего для любого знака выяснить, является он простым или сложным, а если является сложным, то из каких простых он построен.

Теперь можно ввести понятие «смысл» в отличие от понятия «значение» (о понятии смысла см. [1], например). Причем сделаем это без ссылок на понятия «мысль», «сознание», «понятие», «связь понятий» и т. п. Подобные ссылки лишь запутывают дело, ведя к смешению понятий различных наук.

Возможны различные способы определения понятия «смысл», и в частности такой. Смысл простого знака есть его значение. Смысл сложного знака, построенного из данных простых, состоит в том, что он построен из этих простых знаков и все они имеют значение. Отсюда вытекают следующие положения: простой знак имеет смысл, если он имеет значение, и не имеет смысла, если не имеет значения; сложный знак, построенный из данных простых, имеет смысл, если каждый из этих простых знаков имеет значение, и не имеет смысла, если по крайней мере один из этих простых знаков не имеет значения.

Если условиться, что знак имеет значение лишь тогда, когда множество обозначаемых им предметов не пусто, и не имеет, если множество таких предметов пусто, то будет допустима ситуация, когда знак имеет смысл, но не имеет значения (например, «круглый квадрат»). Если исключить из определения смысла то, что смысл простого знака есть его значение, то смысл по договоренности будет свойством только сложного знака (см. [3]). И все эти вариации вполне правомерны.

Таким образом, проблема смысла сводится к исследованию правил построения сложных знаков из простых.

9. Мы, конечно, дали лишь фрагментарный набросок возможной абстрактной теории знаков. По содержанию она в общем виде должна охватить всю ту проблематику теории знаков, которая обычно обсуждается, по форме же должна выступить как применение идей и методов, выработанных современной логикой, к этой проблематике. Так, все содержательные соображения, которые были изложены, могут быть расценены как предисловие или условие интерпретации для некоторой совокупности формальных построений, рассматривающих различного рода фигуры, образуемые из заданных символов.

Необходимо признать, что такого рода теории никогда не отменят необходимости конкретного исследования каждой данной системы знаков. Они могут дать лишь некоторые ориентиры в отношении направлений, способов и систематизации этого конкретного исследования, т. е. всегда будут теориями лишь логико-методологическими.

ЛИТЕРАТУРА

1. Б. В. Бирюков. Теория смысла Г. Фреге. «Применение логики в науке и технике». М., 1960.
2. Л. Ельмслев. Пролегомены к теории языка. «Новое в лингвистике», I, М., 1960.
3. А. А. Зиновьев. Об одном варианте теории определений. «Применение логики в науке и технике», М., 1960.
4. Р. Карнап. Значение и необходимость. М.. 1959.
5. «Новое в лингвистике», II, М., 1962.
6. С. А. Яновская. Предисловие к книге Р. Карнапа [4].

ON THE GROUNDS OF ABSTRACT THEORY OF SIGNS

Summary

The paper deals with the general conditions of constructing abstract (formal) theory (or set of theories) of signs in which the hypothetical arbitrarely assumed properties, relations and associations of any sings independing of their physical nature and practical application are studied.

The basic concepts of the theory of signs («sign», «meaning», «sense» etc.) and possible variations of their definitions are considered. An objective approach to the problems of constructing theory of signs is suggested this excluding unobservable «ideal essences» and preventing the confusion of specific concepts of theory of signs with concepts of psychology, linguistics, philosophy and other sciences.

Ю. К. Лекомцев

ЭЛЕМЕНТЫ ТЕОРИИ ЯЗЫКОВОЙ СОЧЕТАЕМОСТИ

В настоящей работе мы пытаемся наметить круг понятий, пригодных для описания сочетаемости языковых элементов. На исследование данной проблемы автора толкнула неотложная потребность в способе записи, возникавшая всякий раз, когда речь шла о сочетаемости элементов языка, будь то вьетнамские «аналитические» слова в пределах предложения, морфемы в пределах «аналитического» глагола, фонемы в пределах фонологического слога классического тибетского языка.

Понятие сочетаемости (дистрибуция) занимает важное место в практике и теории структурной лингвистики, однако при этом мы очень мало знаем о простейших «пространственных» свойствах элементов языка.

В основном данная работа опирается на наблюдение над наиболее простым объектом — слогом (слогом классического тибетского языка). В разделе о подстановках учитываются некоторые явления синтаксиса (теории предложения).

Неизбежно данная заметка отражает ограниченность исходного материала.

В дальнейшем автор надеется, применяя вводимые здесь понятия при описании все большего числа объектов, обогащать и усложнять теорию.

1. ОСНОВЫ ТЕОРИИ

1.1. Понятия, не получающие определения в рамках теории. Следующий ряд понятий берется нами из общей теории множеств и не получает здесь определения (хотя не все используемые понятия являются неопределяемыми в рамках теории множеств) [1]: множество, элемент множества, отношение, терм отношения, отношение принадлежности ($\in$), соответствие, одно-однозначное соответствие($\longleftrightarrow$), одно-многозначное соответствие ($\leftarrow<$), много-многозначное соответствие ($>-<$), натуральный ряд, число, конечное множество, подмножество, мощность, равномощный.

1. 2. Дано множество M, состоящее из конечного числа элементов, которые мы умеем отождествлять и различать. Среди них пустой элемент (определение будет ниже).

Множество M разбивается на конечные, в принципе, пересекающиеся подмножества элементов, которым поставлены в одно-многозначное соответствие числа натурального ряда от 1 до i: каждому элементу может быть поставлено в соответствие несколько чисел натурального

[1] См. П. С. Александров. Введение в общую теорию множеств и функций. М.—Л., 1948.

ряда, по каждому числу поставлен в соответствие только один элемент. Такие подмножества назовем к о р т е ж а м и (K).

Натуральные числа от 1 до i, поставленные в соответствие элементам M, будем называть и н д е к с а м и.

Каждый индекс вместе с его элементом назовем п а р о й.

Число пар кортежа из M назовем м о щ н о с т ь ю этого кортежа. Кортежи с равными числами пар назовем р а в н о м о щ н ы м и.

Две пары считаются тождественными, если тождественны элементы и равны их индексы.

Два кортежа считаются тождественными, если они равномощны и составляющие их пары тождественны.

Нас интересуют подмножества равномощных кортежей.

Обозначим мощность символом ω.

$\omega(K_h) = m$ или $\underset{m}{K_h}$ будет обозначать кортеж мощности m.

В настоящем разделе работы нас интересует некоторое множество H_i кортежей $\underset{m}{K}$ фиксированной мощности.

Введем понятие п о д к о р т е ж а K^1_k кортежа K_k: назовем подкортежем конечное множество пар мощности, не превышающей мощности K_k, состоящее только из таких пар, которые принадлежат K_k.

$\omega(K^1_k) = B$ будет означать: «мощность подкортежа K^1_k равна B», $\underset{B}{K^1_k}$ будет означать: «подкортеж K^1_k мощности B». Любой кортеж $\underset{m}{K_g}$ может рассматриваться как подкортеж $\underset{m}{K^1_g}$.

О т н о ш е н и е R_1

Два элемента a и b, принадлежащие M, связаны отношением R_1, если a и b принадлежат одному и тому же кортежу K_n.

$(\forall a \in M)\ (\forall b \in M)\ (a \in K_n)\ \&\ (b \in K_n) \equiv aR_1b$ [2].

О т н о ш е н и е P

Два элемента a и b из M связаны отношением P, если они имеют индексы, отличающиеся на 1.

$(\forall a \in M)\ (\forall b \in M)\ (a \leftarrow< \text{Index } m)\ \&\ (b \leftarrow< \text{Index } n)\ \&\ (\overline{(\text{Index } m - \text{Index } n = 1) \equiv (\text{Index } n - \text{Index } m = 1)}) \equiv aPb.$

О т н о ш е н и е R_0

Два элемента a и b из M связаны отношением R_0, если они принадлежат одному кортежу и их индексы отличаются на 1.

$(\forall a \in M)\ (\forall b \in M)\ (a \in K_n)\ \&\ (b \in K_n)\ \&\ (a \leftarrow< \text{Index } m)\ \&\ (b \leftarrow< \text{Index } n)\ \&$
$\&\ (\overline{(\text{Index } n - \text{Index } m = 1) \equiv (\text{Index } m - \text{Index } n = 1)}) \equiv aR_0b$ или
$(\forall a \in M)\ (\forall b \in M)\ (aR_1b)\ \&\ (aPb) \equiv aR_0b.$

Определим отношения R_1, P и R_0 для элемента и подкортежа.

О т н о ш е н и е R_1

Элемент a и подкортеж K^1_k из M связаны отношением R_1, если они принадлежат одному и тому же кортежу K_l, т. е. K^1_k имеет мощность $m - 1$, а элемент a имеет индекс (из числа от 1 до i), отсутствующий в K^1_k.

$(\forall a \in M)\ (\forall K^1_k \in M)\ (a \in K_l)\ \&\ (K^1_k \in K_l) \equiv aR_1K^1_k$
или $(\forall a \in m)\ (\forall K^1_k \in M)\ (\omega(K^1_k) = m - 1)\ \&\ (\exists \text{ Index } j)\ \&\ (\text{Index } j \rightarrow> a)\ \&$
$\&\ (\text{Index } j \overline{\in} \text{ Мн. Index } (K^1_k)) \equiv aR_1K^1_k.$

О т н о ш е н и е P

Элемент a и подкортеж K^1_k из M связаны отношением P, если индекс a отличается на 1 от одного из индексов K^1_k.

---

[2] Значения логических символов: $\forall$ — квантор всеобщности; $\exists$ — квантор существования, $\&$ — конъюнкция, $\leftarrow$ — импликация, $\equiv$ — эквивалентность, $\bar{a}$ — отрицание a и т. д. См. Д. Г и л ь б е р т, В. А к к е р м а н. Основы теоретической логики. М., 1947.

$(\forall a \in M)\ (\forall K_k^1 \in M)\ (\exists \text{ Index } j)\ \&\ (\text{Index } j \rightarrow a)\ \&\ (\exists \text{ Index } g)\ \&$
$\&\ (\text{Index } g \in \text{Мн. Index } (K_k^1))\ \&\ (\overline{(g-j=1) \equiv (j-g=1)}) \equiv aPK_k^1.$

Отношение R_0

Элемент a и подкортеж K_k^1 из M связаны отношением R_0, если a и K_k^1 принадлежат одному и тому же кортежу K_l и если индекс a отличается на 1 по крайней мере от одного из индексов, принадлежащих K_k^1.

$(\forall a \in M)\ (\forall K_k^1 \in M)\ (aR_1K_k^1)\ \&\ (aPK_k^1) \equiv aR_0K_k^1.$

Отношение G

Элементы a и c из M связаны отношением G, если a и b имеют один и тот же индекс:

$(\forall a \in M)\ (\forall c \in M)\ (a \leftarrow < \text{Index } m)\ \&\ (c \leftarrow < \text{Index } m) \equiv aGc.$

Из определения следует, что либо aR_1c и $a=c$, либо $\overline{aR_1c}$.

Отношение R_2

Элементы a и c из M связаны отношением R_2, если они не принадлежат одному и тому же кортежу, имеют равные индексы и образуют кортеж с тождественными подкортежами мощности $m-1$.

$(\forall a \in M)\ (\forall c \in M)\ (\overline{aR_1c})\ \&\ (aGc)\ \&\ (aR_{0\ m-1}K_g^1)\ \&\ (cR_{0\ m-1}K_g^1) \equiv aR_2c.$

Определим отношения, введенные нами ранее для элементов и элемента с подкортежем, также и для подкортежей.

Предварительно дадим несколько определений, относящихся к понятиям нуля и нулевого подкортежа.

Введем понятие пустого элемента (нуля) — 0. Нулем (пустым элементом) назовем такой элемент из M, что

$(\forall a \in M)\ (\exists o \in M)\ (aR_0o) \equiv a.$

Нулевым подкортежем назовем подкортеж мощности B — $K^1_{B_0} =$

$= (o_1R_0o_2R_0 \ \ldots\ R_0o_b)$

по определению нуля

$K^1_{B_0} = o_1R_0o_2R_0 \ \ldots\ R_0o_b = 0.$

Множество, состоящее из пустого элемента, имеет мощность, равную нулю.

Следовательно, $K^1_{B_0} = K^1_0$.

Отношение R_1

Два подкортежа K_k^1 и K_l^1 из M связаны отношением R_1, если они принадлежат одному кортежу, т. е. если $\omega(K_k^1) + \omega(K_l^1) = m$, и если множество индексов, принадлежащих K_k^1 или K_l^1, включает все индексы от 1 до m:

$$(\forall K_k^1 \in M)\ (\forall K_l^1 \in M)\ (K_k^1 \in K_m)\ \&\ (K_l^1 \in K_m) \equiv K_k^1R_1K_l^1$$

или с дополнительным условием

$(\forall K_k^1 \in M)\ (\forall K_l^1 \in M)\ (\omega(K_k^1) + \omega(K_l^1) = m)\ \&\ (\text{Мн. Index } (K_k^1 \cup K_l^1) =$
$= (1, 2, 3, \ \ldots, m))$[3]. 7.

Отношение P

Наибольший индекс подкортежа K_k^1 обозначим $\text{Index}_{\max}(K_k^1)$. Наименьший индекс подкортежа K_k^1 обозначим $\text{Index}_{\min}(K_k^1)$.

Два подкортежа K_k^1 и K_l^1 из M связаны отношением P, если $\text{Index}_{\max}(K_k^1) = m$ и $\text{Index}_{\min}(K_l^1) = n$ отличаются на 1.

---

[3] Мн. Index $(K_k^1 \cup K_l^1)$ означает «множество индексов, принадлежащих хотя бы одному из подкортежей».

$(\forall K_k^1 \in M)\ (\forall K_l^1 \in M)\ (\mathrm{Index}_{\max}(K_k^1) = m)\ \&\ \mathrm{Index}_{\min}(K_l^1) = n)\ \&\ (n - m = 1) \equiv K_k^1 P K_l^1$.

Отношение R_0

Два подкортежа K_k^1 и K_l^1 из M связаны отношением R_0, если они принадлежат одному кортежу K_m и если $\mathrm{Index}_{\max}(K_k^1)$ и $\mathrm{Index}_{\min}(K_l^1)$ отличаются на 1.

$(\forall K_k^1 \in M)\ (\forall K_l^1 \in M)\ (K_k^1 R_1 K_l^1)\ \&\ (K_k^1 P K_l^1) \equiv K_k^1 R_0 K_l^1$.

Отношение G

Два подкортежа K_k^1 и K_l^1 из M связаны отношением G, если они равномощны и имеют попарно равные индексы.

$$(\forall K_k^1 \in M)(\forall K_l^1 \in M)(\omega(K_k^1) = \omega(K_l^1))\ \&\ (\mathrm{Index}\, i\,(K_k^1) = \mathrm{Index}\, i\,(K_l^1)) \equiv K_k^1 G K_l^1.$$
$$i = h,\ h+1,\ h+2, \ldots \quad i = h,\ h+1,\ h+2, \ldots$$

Отношение R_2

Два подкортежа K_k^1 и K_l^1 из M связаны отношением R_2, если они не принадлежат одному кортежу, равномощны, имеют попарно равные индексы и связаны отношением R_0 с одним и тем же подкортежем K_p^1, причем мощность $K_k^1 R_0 K_p^1 = m$ (соответственно, мощность $K_k^1 R_0 K_p^1 = m$).

$$(\forall K_k^1 \in M)\ (\forall K_l^1 \in M)\ (\forall K_p^1 \in M)\ (K_k^1 R_1 K_l^1)\ \&\ (K_k^1 G K_t^1)\ \&\ (K_k^1 G K_l^1)\ \&$$
$$\&\ (K_k^1 R_0 K_p^1)\ \&\ (K_l^1 R_0 K_p^1)\ \&\ (\omega(K_k^1, K_l^1) + \omega(K_p^1) = m) \equiv K_k^1 R_2 K_l^1.$$

Частным случаем является тот, когда $\omega(K_k^1, K_l^1) = m$, т. е. когда мы имеем дело с кортежами.

Тогда

$$(\forall \underset{m}{K_k^1} \in M)\ (\forall \underset{m}{K_k^1} \in M)\ (\forall \underset{o}{K_p^1} \in M)\ (\overline{K_k^1 R_1 K_l^1})\ \&\ (\underset{m}{K_k^1} G \underset{m}{K_l^1})\ \&\ (\underset{m}{K_k^1} R_0 \underset{o}{K_p^1})\ \&$$
$$\&\ (\underset{m}{K_l^1} R_0 \underset{o}{K_p^1})\ \&\ (\omega(\underset{m}{K_k^1}, \underset{m}{K_l^1}) + \omega(\underset{o}{K_p^1}) = m),$$

т. е. все кортежи из H связаны отношением R_2.

Наиболее важны для нашего исчисления отношения R_0 и R_2.

Отношения R_0 и R_2 обладают следующими характеристиками:

Отношение	Рефлексивность	Симметричность	Транзитивность
R_0	—	+	—
R_2	+	+	+

Условные знаки: + означает прямое отношение; — означает антиотношение.

Из множества H_i кортежей мощности m выделим множество J_i кортежей, удовлетворяющих следующему условию:

(1) элементы в J_i связаны отношениями R_0 и R_2,

(2) все элементы J_i, имеющие один и тот же индекс, связаны отношением R_2.

Множество элементов с одним и тем же индексом, являющихся термами по R_2, назовем классом (Kl):

$(\forall a_i, b_i, \ldots m_i^1)\ (R_2\ (a_i, \ldots b_i, \ldots m_i^1)) \equiv \text{Класс}\ (Kl)$.

Конечное множество классов, которым поставлены во взаимно-однозначное соответствие числа натурального ряда от 1 до i (индексы), назовем суперкортежем (SKO).

Индекс с его классом назовем парой суперкортежа. Число пар, входящих в суперкортеж, называется его мощностью. Условно обозначим: $\omega(SKO)=n$ или $\overset{n}{SKO}$.

Определим отношение R_0 для классов A и B: классы A и B связаны отношением R_0, если они принадлежат одному и тому же суперкортежу, а их индексы отличаются на 1.

$$(\forall A \in J)\ (\forall B \in J)\ (A \in \overset{n}{SCO}) \& (B \in \overset{n}{SCO}) \& \overline{(\text{Index } (A) - \text{Index } (B) = 1) \equiv} \equiv \overline{(\text{Index } (B) - \text{Index } (A) = 1)} \equiv AR_0B$$

Таким образом, подмножеству кортежей множества J_i будет соответствовать один суперкортеж.

Например, кортежам $(a_1R_0C_2)\ R_2\ (b_1R_0c_2)\ R_2\ (a_1R_0d_2)\ R_2\ (b_1R_0d_2)$ соответствует суперкортеж $(a_1R_2b_1)\ R_0\ (C_2R_2d_2)_2$.

Множество, включающее наряду с множеством J_i кортежей также суперкортеж, назовем множеством J_i^1.

В множестве J_i^1 определены два отношения, одно из которых R_2 связывает индексированные элементы и кортежи, другое R_0 — индексированные элементы и их классы.

Вернемся к характеристикам отношений R_0 и R_2. Для отношений R_0 и R_2 в множестве J^1 постулируется следующее.

1. Коммутативность R_0:
   $aR_0b = bR_0a$;
2. Коммутативность R_2:
   $aR_2c = cR_2a$;
3. Ассоциативность R_0:
   $(aR_0b)\,R_0c = aR_0\,(bR_0c)$;
4. Ассоциативность R_2:
   $(aR_2b)\,R_2c = aR_2\,(bR_2c)$;
5. Дистрибутивность R_0кR_2:
   $aR_0(bR_2c = (aR_0b)\,R_2\,(aR_0c)$.

Дистрибутивность R_0 и R_2, чрезвычайно важная для нас в дальнейшем, проявляется в двух противоположных операциях: в размещении терма по R_0 по отношению каждого из термов по R_2 в пределах скобки (прямая операция) и в вынесении за скобку терма по R_0, связанного с каждым термом по R_2.

2. СПОСОБЫ ЗАПИСИ СТРУКТУР

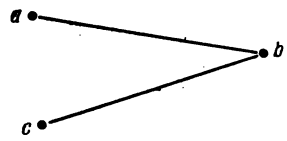
С настоящего момента в целях удобства записи мы введем специальные символы для отношений R_0 и R_2.

Вместо aR_0b мы вводим $a \cdot b$ и вместо aR_2c мы вводим $a\,//\,c$.

С целью обозначения понимания записей типа $a \cdot b\,//\,c$ мы вводим скобки: $(a \cdot b)\,//\,c$ и $a \cdot (b\,//\,c)$ — различные структуры. Запись, использующую символы $\cdot$, $//$, ( ), мы условно будем называть алгебраической.

Те же отношения R_0 и R_2 могут быть представлены в геометрической записи: aR_0b может быть изображено следующим образом $a \cdot \text{———} \cdot b$. Точки обозначают здесь термы отношения R_0, линия—отношение R_0, aR_2c изображается $\begin{matrix} a \\ \vdots \\ c \end{matrix}$, где отсутствие линии между точками указывает на отношение R_2, а параллельное расположение точек указывает на совпадение индексов у a и c.

Структура, например $(a\,//\,c)\cdot b$, записывается следующим образом:



Любую фигуру из точек (вершин) и линий (неориентированных ребер) будем называть графом.

В дальнейшем, кроме точек и стрелок, в состав графа нам придется ввести ряд других элементов (например, аналог скобок). Основная запись, которой мы будем пользоваться, — алгебраическая[4].

В существенных местах параллельно будет даваться и геометрическая запись.

3. РНДС И ЕЕ КОРРЕЛЯТ

Назовем суперкортеж из множества J_i^1 равномерной нормальной дистрибутивной структурой (РНДС) и назовем множество кортежей J_1 коррелятом РНДС.

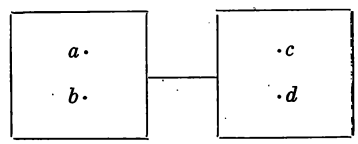
В нашей алгебраической записи РНДС и коррелят для четырех кортежей с $\omega(K_g)=2$ имеют следующий вид:

РНДС: $(a\,//\,b)\cdot(c\,//\,d)$;

коррелят: $(a\cdot c)\,//\,(b\cdot c)\,//\,(a\cdot d)\,//\,(b\cdot d)$.

Обе формулы связаны законом дистрибутивности: РНДС преобразуется в свой коррелят путем распределения терма по R_0 $(\cdot)$ относительно термов по R_2 $(//)$; коррелят преобразуется в РНДС путем вынесения за скобки общего терма по R_0 $(\cdot)$.

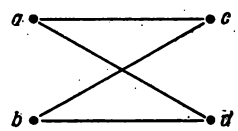
В терминах геометрической записи имеем



(1)

для РНДС

и



(2)

для ее коррелята. Граф вида (1) мы назовем укрупненным графом. Рамка соответствует здесь скобкам, в которые в РНДС берется сложный терм по отношению R_0:

Любая РНДС изображается укрупненным графом.

---

[4] Мы называем здесь запись алгебраической условно, так как R_0 и R_2 суть бинарные отношения, а операции алгебры суть тернарные отношения (например, сложение или умножение). Но R_0 и R_2 связаны законом дистрибутивности так же, как умножение и сложение (т. е. односторонне: первое отношение дистрибутивно ко второму, но не наоборот).

Граф из рамок называется определенным графом (сокращенно — просто графом).

В следующем разделе мы займемся соотношениями между РНДС и ее коррелятом.

4. СООТНОШЕНИЕ МЕЖДУ РНДС И ЕЕ КОРРЕЛЯТОМ

(1) Набор символов:

λ — В-длина РНДС [5]
l — В-длина скобки РНДС
m — число пар скобок РНДС
p — число символов // в РНДС
r — число символов · в РНДС
L — общая длина РНДС
Θ — экономность записи РНДС

λ_1 — В-длина коррелята
l_1 — В-длина скобки коррелята
m_1 — число пар скобок коррелята
p_1 — число символов // коррелята
r_1 — число символов · коррелята
L_1 — общая длина коррелята

РНДС и ее коррелят связаны следующими соотношениями (слева мы выражаем понятия, связанные с коррелятом, в понятиях, связанных в РНДС, справа — наоборот).

(2) Соотношение	Обратное соотношение
Соотношение 1	Соотношение 1′
$l_1 = m$	$m = l_1$
Соотношение 2	Соотношение 2′
$m_1 = l^m$	$l^m = m_1 = l^{l_1} = m_1$
Соотношение 3	Соотношение 3′
$\lambda = lm$	$l = \sqrt[l_1]{m_1}$
	$\lambda = l_1 \cdot \sqrt[l_1]{m_1}$
Соотношение 4	Соотношение 4′
$\lambda_1 = l^m m$	$\lambda_1 = l_1 \cdot \sqrt[2l_1]{m_1}$
Соотношение 5	Соотношение 5′
$p = (l-1) \cdot m$	$p = l_1 \cdot \left(\sqrt[l_1]{m_1} - 1\right)$
Соотношение 6	Соотношение 6′
$r = m - 1$	$r = l_1 - 1$
Соотношение 7	Соотношение 7′
$p_1 = l^m - 1$	$(p_1 = m_1 - 1;\ m_1 = l^m)$
	$p_1 = \sqrt[2l_1]{m_1} - 1$
Соотношение 8	Соотношение 8′
$r_1 = l^m (m-1)$	$r_1 = (l_1 - 1) \cdot \sqrt[2l_1]{m_1}$

Соотношения 1—8 доказываются по индукции. Приведем лишь доказательство соотношения 1:

$$l_1 = m.$$

Соотношение доказывается индукцией по числу скобок в РНДС. Введем величину $\Lambda_m = m - 1$.

(A) Пусть l (b — длина скобки РНДС) = const = 2.

1. Соотношение справедливо для $\Lambda_m = 1$ (т. е. $m = 2$)

$$(a // b) \cdot (c // d) = (a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d).$$

2. Предположим, что оно справедливо для $\Lambda_m = n$ $(m = n + 1)$.

[5] *B*-длиной мы называем число буквенных символов.

3. Докажем, что добавление одной скобки в РНДС не нарушит соотношения $l_1 = m$.

Имеется $n+1$ скобок в РНДС и некоторое число K скобок с $l_1 = n+1$ в корреляте.

Добавление новой скобки в РНДС означает включение в каждую скобку коррелята по одному буквенному символу (терму отношения // // в РНДС) с одновременным увеличением числа скобок коррелята вдвое ($2K$), так как $l = 2$. Таким образом, $n+2$ скобкам РНДС соответствует $l_1 = n+2$ и для $Л_m = n+1$; $m = l_1$.

(В) При переменном l соответствие $m = l_1$ не нарушится, так как термы по // из одной скобки РНДС не могут входить в одну скобку коррелята: это противоречило бы определению отношения $R_2(//)$, согласно которому термы по // не принадлежат тому же подмножеству Sq_k. Увеличение l ведет к увеличению числа скобок коррелята, но не к увеличению l_1.

Определение. Общей длиной РНДС (L) назовем сумму буквенной длины РНДС (λ), числа символов // (p), числа символов $\cdot$ (r) и числа пар скобок (m).

$$L = \lambda + p + r + m$$

или

$$L = m(2l - 1) + 2m - 1,$$

или

$$L = l_1(2\sqrt[l_1]{m_1} + l_1 - 1) - 1.$$

Определение. Общей длиной коррелята (L_1) назовем сумму буквенной длины (λ_1), числа символов // (p_1), числа символов $\cdot$ (r_1) и числа пар скобок (m_1).

$$L_1 = \lambda_1 + p_1 + r_1 + m_1$$

или

$$L_1 = l^m(2m + 1) - 1,$$

или

$$L_1 = 2l_1\sqrt[2l_1]{m_1} + m_1 - 1.$$

Определение. Экономностью записи РНДС ($Э$) назовем отношение $Э = \frac{L_1}{L}$.

$$Э = \frac{L_1}{L} = \frac{l^m(2m+1) - 1}{m(2l + m - 1) - 1},$$

$$Э = \frac{L_1}{L} = \frac{2l_1\sqrt[2l_1]{m_1} + m_1 - 1}{l_1\left(2\sqrt[l_1]{m_1} + l_1 - 1\right) - 1}.$$

Формулу $\frac{l^m(2+1) - 1}{m(2l + m - 1) - 1}$, выражающую экономность записи вида РНДС по сравнению с записью вида коррелята РНДС, можно представить на графике (см. график на стр. 31) как функцию от целочисленных аргументов l и m:

$$f(l, m) = \frac{l^m(2m+1) - 1}{m(2l + m - 1) - 1}.$$

Чтобы получить график на плоскости, возьмем $l = m$. Для нас имеют смысл значения аргументов $\geqslant 2$, однако на графике мы отразим более широкую область значений аргументов, введя значение 1, —1 и 0 (см. табл. 1).

Остановимся на еще одном важном соотношении между РНД ее коррелятом.

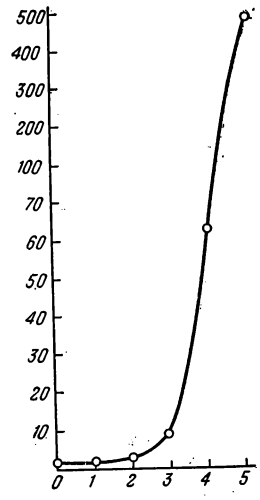
Таблица 1

Значения $l=m$	—5	—4	—3	—2	—1	0	1	2	3	4	5
Значения $f(l, m)$	—0,01	—0,02	—0,03	—0,13	0	1	2	2,1	8,17	63,5	498

Пусть N — число различных элементов, входящих в коррелят, f — число вхождений каждого из N элементов в коррелят. Тогда справедливо следующее соотношение: $f=l^{m-1}$ или $f=m_1\frac{l_1-1}{l_1}$. Второй вариант соотношения выводится из первого. Действительно,

$$l^{m-1}=\frac{l^m}{l};\ \frac{l^m}{l}=\frac{m_1}{l};\ l=\sqrt[l_1]{m_1};$$

$$\frac{l^m}{l}=\frac{m^1}{\sqrt[l_1]{m_1}}=m_1\cdot m_1^{\frac{-1}{l_1}}=m_1^{1-\frac{1}{l_1}}=m_1^{\frac{l_1-1}{l_1}}.$$



Докажем это соотношение. Пусть f есть число вхождений каждого элемента РНДС по отношению // в коррелят $f=l^{m-1}$.

Докажем это положение по индукции.

(I) Сначала возьмем постоянное $l=2$. Прежде всего выберем исходное отношение; это будет $Л_m$— длина по числу скобок.

Коррелят «наименьшей» длины ($Л_m$) типа $(a\cdot c)\,//\,(b\cdot c)\,//\,(a\cdot d)\,//\,(b\cdot d)$ имеет РНДС «наименьшей» длины $(a\,//\,b)\cdot(c\,//\,d)$. Условимся единицей длины ($Л_m$) РНДС считать $m=2$, т. е. $Л_m$ РНДС $=m-1$.

1. Базис индукции. Покажем, что соотношение выполняется для коррелята РНДС с $Л_m=1$ действительно:

$$f(a)=2=2^1=l^{m-1};\ f(c)=2=2^1=l^{m-1}.$$
$$f(b)=2=2^1=l^{m-1};\ f(d)=2=2^1=l^{m-1}.$$

2. Допустим, что соотношение справедливо для коррелята РНДС с $Л_m=n$.

3. Индукционный шаг. Докажем, что оно справедливо для $Л_{m=n+1}$. При $Л=n$, $m=n+1$, а $m_1=l^m=l^{n+1}$, при добавлении к РНДС одной скобки $m=n+2$. Каждый элемент новой скобки распределяется по l^{n+1} скобок, т. е. каждый новый элемент входит в коррелят l^{n+1} раз при $m=n+2$, т. е. входит l^{m-1} раз. Было допущено, что каждый из «старых» элементов входит в коррелят РНДС длины n l^{m-1} раз при $m=n+1$, т. е. входит n раз.

При добавлении новой скобки к РНДС ($Л_m=n+1$, $m=n+2$) число вхождений каждого из старых элементов увеличилось в l раз, т. е. $=l^n\cdot l=l^{n+1}$.

Таким образом, в корреляте РНДС с $Л_m=n+1$ каждый элемент содержится l^{m-1} раз.

(II) Докажем то же для переменного l при постоянном $m=2$. Возьмем исходное отношение $Л_1$, т. е. число букв в скобке РНДС, и положим, что при $l=2$ $Л_1=1$.

1. Базис индукции. Положение справедливо при $Л_1=1$ (см. стр. 29, 1).

2. Допустим, что оно справедливо при $Л_l=n$; $l=n+1$.

3. Индукционный шаг. Возьмем $Д_1 = n + 1$, это значит, что мы увеличили число элементов скобки РНДС по // на один.

$$l^1 = n + 2.$$

Распределяя отношение, мы связываем каждый новый элемент скобки РНДС с l «старыми» элементами и $m - 1$ «новыми» элементами, т. е. число вхождений «нового» элемента

$$f = l + m - 1 = n + 1 + m - 1 = n + 2,$$

каждый «старый» элемент сочетается с l «старыми» элементами и $m - 1$ «новыми»: $f = l + m - 1 = n + 2$, таким образом $f = l^1 = l^{m-1}$. Соотношение доказано. Соотношение $f = m_1^{\frac{l_1 - 1}{l_1}}$ является критерием того, что данный набор кортежей есть коррелят РНДС.

5. ДИСТРИБУТИВНЫЕ СТРУКТУРЫ

5.1. Понятие РНДС и коррелята интересуют нас постольку, поскольку от них можно сделать дальнейший шаг, приближающий нас к законам сочетаемости реальных языковых элементов. Введем новое понятие — понятие дистрибутивной структуры (ДС).

Под дистрибутивной структурой мы будем понимать любое подмножество элементов из J, в котором представлены оба отношения R_0 и R_2 $(\cdot\ ,\ //)$.

Дистрибутивные структуры могут быть получены из суперкортежа или множества кортежей J благодаря операции подстановки, т. е. благодаря автоморфному отображению особого вида множества J^1.

Дистрибутивные структуры, полученные применением операции подстановки к суперкортежам (т. е. РНДС), так же, как и сами РНДС, являющиеся результатом нулевой операции, назовем прямыми дистрибутивными структурами (ПДС).

Дистрибутивные структуры, полученные применением операции подстановки к множеству J (т. е. к корреляту РНДС), а также сами корреляты РНДС назовем обратными дистрибутивными структурами (ОДС).

5.2. Прямые дистрибутивные структуры. Элементом РНДС (Э) назовем терм по отношению $R_2(//)$, а членом РНДС (Ч) назовем терм (сложный терм) по отношению R_0. Саму структуру (РНДС) сокращенно обозначим через С. РНДС, в которую производится подстановка, назовем основной РНДС; РНДС, из которой производится подстановка, назовем исходной; ПДС, возникшая в результате подстановки, будет названа результирующей ПДС. Операция подстановки может быть разбита на три типа.

1) Подстановка, при которой результирующая ПДС содержит большее число символов, чем основная РНДС, называется расширением.

2) Подстановка, при которой результирующая ПДС содержит меньшее число символов, чем основная РНДС, называется сжатием.

3) Подстановка, при которой результирующая ПДС содержит то же число символов, что и основная РНДС, называется обменом. При расширении возможна подстановка члена вместо элемента, структуры вместо элемента и структуры вместо члена (расширение-ЭЧ, расширение-ЭС, расширение-ЧС).

При сжатии возможна подстановка члена вместо структуры, элемента вместо структуры, элемента вместо члена (сжатие-СЧ, сжатие-СЭ, сжатие-ЧЭ).

При обмене возможна подстановка элемента вместо элемента, члена вместо члена, структуры вместо структуры (обмен-ЭЭ, обмен-ЧЧ, обмен-СС).

Основные типы подстановок могут быть представлены в табл. 2.

Таблица 2

Основная РНДС \ Исходная РНДС	Элемент	Член	Структура
Элемент	1. Обмен-ЭЭ	2. Расширение-ЭЧ	3. Расширение-ЭС
Член	4. Сжатие-ЧЭ	5. Обмен-ЧЧ	6. Расширение-ЧС
Структура	7. Сжатие-СЭ	8. Сжатие-СЧ	9. Обмен-СС

Примеры:

Пусть $(a \,//\, b) \cdot (c \,//\, d)$ — основная, а $(a_1 \,//\, b_1) \cdot (c_1 \,//\, d_1)$ — исходная РНДС. Тогда все случаи таблицы можно проиллюстрировать следующим образом (случай 1 в примере не нуждается).

2. Расширение-ЭЧ:

$$(a_1 \,//\, b_1)_a \,//\, b) \cdot (c \,//\, d).$$

3. Расширение-ЭС:

$$(((a_1 \,//\, b_1) \cdot (c_1 \,//\, d))_a \,//\, b) \cdot (c \,//\, d).$$

6. Расширение-ЧС:

$$((a_1 \,//\, b_1) \cdot (c_1 \,//\, d_1))_{(a \,//\, b)} \cdot (c \,//\, d).$$

4. Сжатие-ЧЭ:

$$(a_1) \cdot (c \,//\, d).$$

7. Сжатие-СЭ [6]:

$$(a_1)_{((a \,//\, b) \,\cdot\, (c \,//\, d))} \cdot$$

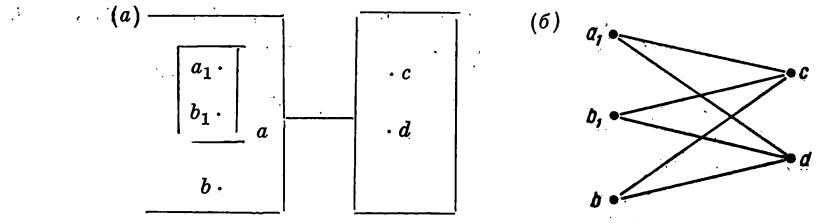
8. Сжатие-СЧ:

$$(a_1 \,//\, b_1)_{((a \,//\, b) \,\cdot\, (c \,//\, d))} \cdot$$

Эти же примеры представим в виде графов.

2. Расширение-ЭЧ

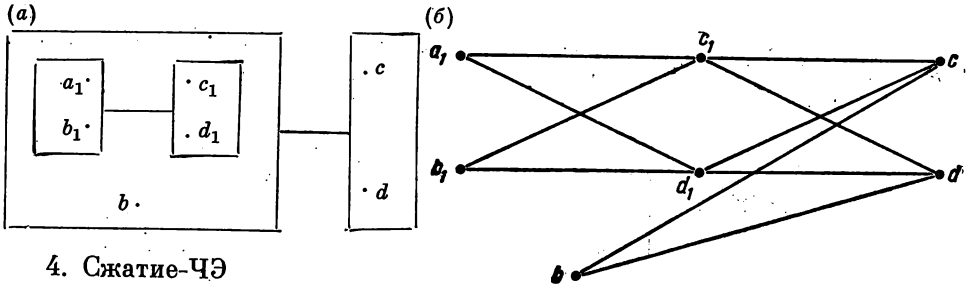
$((a_1 \,//\, b_1)_a \,//\, b) \cdot (c \,//\, d) = (a_1 \,//\, b_1 \,//\, b) \cdot (c \,//\, d) = (a_1 \cdot c) \,//\, (b_1 \cdot c) \,//\, (b \cdot c) \,//\, (a_1 \cdot d) \,//\, //\, (b_1 \cdot d) \,//\, (b \cdot d)$



3. Расширение-ЭС

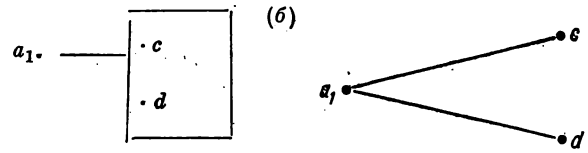
$(((a_1 \,//\, b_1) \cdot (c_1 \,//\, d_1))_a \,//\, b) \cdot (c \,//\, d) = [(a_1 \cdot c_1) \,//\, (b_1 \cdot c_1) \,//\, (a_1 \cdot d_1) \,//\, (b_1 \cdot d_1) \,//\, b] \cdot$
$\cdot (c \,//\, d) = (a_1 \cdot c_1 \cdot c) \,//\, (b_1 \cdot c_1 \cdot c) \,//\, (a_1 \cdot d_1 \cdot c) \,//\, (b_1 \cdot d_1 \cdot c) \,//\, (b \cdot c) \,//\, (a_1 \cdot c_1 \cdot d) \,//$
$//\, (b_1 \cdot c_1 \cdot d) \,//\, (a_1 \cdot d_1 \cdot d) \,//\, (b_1 \cdot d_1 \cdot d) \,//\, (b \cdot d)$

[6] Подстановки 7 и 8 приводят к результирующим образованиям, не являющимся по определению ДС. Мы будем считать их действительными лишь в случае, когда основная РНДС есть часть другой РНДС.



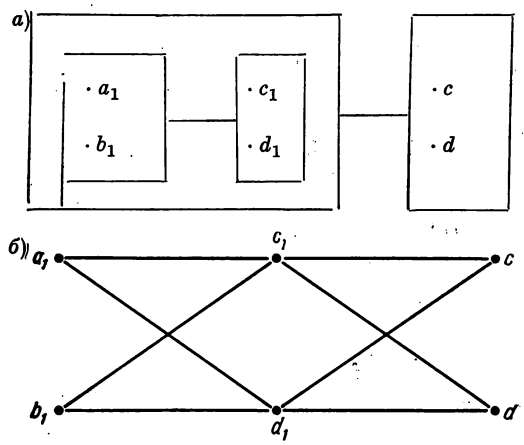
4. Сжатие-ЧЭ

$$(a_1)_{(a\,//\,b)} \cdot (c\,//\,d) = (a_1 \cdot c)\,//\,(a_1 \cdot d)$$



6. Расширение-ЧС

$((a_1\,//\,b_1) \cdot (c_1\,//\,d_1))_{(a\,//\,b)} \cdot (c\,//\,d) = (a_1//\,b_1) \cdot (c_1\,//\,d_1) \cdot (c//\,d) = (a_1 \cdot c_1\,//\,b_1 \cdot c_1//$
$//a_1 \cdot d_1\,//\,b_1 \cdot d_1) \cdot (c\,//\,d) = (a_1 \cdot c_1 \cdot c)\,//\,(b_1 \cdot c_1 \cdot c)\,//\,(a_1 \cdot d_1 \cdot c)\,//\,(b_1 \cdot d_1 \cdot c)//\,(a_1 \cdot$
$\cdot c_1 \cdot d)\,//\,(b_1 \cdot c_1 \cdot d)\,//\,(a_1 \cdot d_1 \cdot d)\,//\,(b_1 \cdot d_1 \cdot d)$



7. Сжатие-СЭ

$\cdot a$

8. Сжатие-СЧ

b

1. Обмен-ЭЭ

$((a_1)_a\,//\,b) \cdot (c\,//\,d)$

5. Обмен-ЧЧ

$(a_1\,//\,b_1)_{a\,//\,b} \cdot (c\,//\,d)$

9. Обмен-СС

$((a_1\,//\,b_1) \cdot (c_1\,//\,d_1))_{(a\,//\,b)\,\cdot\,(c\,//\,d)}$

По характеру основной и исходной РНДС мы будем различать два типа подстановок: автоморфные подстановки, когда основная и исходная РНДС совпадают, и гетероморфные подстановки, когда основная и исходная РНДС различны.

Операция подстановки может быть применена к основной РНДС более одного раза. В таких случаях нам следует различать одновременное и последовательное проведение нескольких подстановок.

Поставим в соответствие множеству операций упорядоченное множество элементов, которые мы будем называть моментами и обозначать t_n; $n = 1, 2, 3, \ldots$

Каждый момент t_k определяет такт подстановки. Подстановки, имеющие один и тот же момент, относятся к одному и тому же такту. ПДС, полученная в результате n-тактной подстановки, называется n-тактной ПДС. Две или несколько операций подстановки над одной основой РНДС, для которых безразличен порядок выполнения, получают одинаковые моменты (и называются однотактными), например (расширение-ЭЧ)$_{T_1}$ и (расширение-ЧС)$_{T_1}$:

основная РНДС: $(a\,//\,b) \cdot (c\,//\,d)$;

исходная РНДС: $(a_1 // b_1) \cdot (c_1 // d_1)$

$$\frac{\begin{array}{l} a = (a_1 // b_1) \\ (c // d) = (a_1 // b_1) \cdot (c_1 // d_1) \end{array}}{((a_1 // b_1)_a // b) \cdot ((a_1 // b_1) \cdot (c_1 // d_1))_{(c // d)}}.$$

Две или несколько операций подстановки над одной РНДС, совершаемые последовательно, получают разные моменты (и называются двух- или n-тактными), например (сжатие-ЧЭ)$_{T_1}$ и (сжатие-СЧ)$_{T_2}$:

основная РНДС: $(a // b) \cdot (c // d)$;
исходная РНДС: $(a_1 // b_1) \cdot (c_1 // d_1)$

$$\frac{(a // b) = a_1}{(a_1)_{(a // b)} \cdot (c // d)} T_1$$

$$\frac{a_1 \cdot (c // d) = (c_1 // d_1)}{(c // d)_{(a_1)_{(a // b)}} \cdot (c // d)} T_2.$$

В том случае, когда один и тот же тип однотактной подстановки применяется K раз, мы будем писать, например: K (расширение-ЭЧ)$_{T_1}$, т. е. в данном примере K элементов основной РНДС одновременно заменяются членами исходной РНДС.

Если K равно общему числу заменяемых единиц (в данном примере общему числу элементов основной РНДС), то мы пишем, опять применительно к нашему примеру:

$(\forall$ Э) (расширение-ЭЧ)$_{T_1}$.

Из всевозможных результирующих ПДС мы выделим несколько типов ПДС, особенно важных при применении теории.

Трансференция. В результате применения операции обмена возникает структура, которую мы называем трансференцией[7].

Примеры:

$$\frac{\begin{array}{l} (a // b) \cdot (c // d) \\ (a // b) \cdot (c // d) \\ a = c \end{array}}{((c)_a // b) \cdot (c // d)}$$

или

$$\frac{\begin{array}{l} (a // b) \cdot (c // d) \\ (a_1 // b_1) \cdot (c_1 // d_1) \\ a // b = c_1 // d_1 \end{array}}{(c_1 // d_1) \cdot (c // d)}$$

Цикл$_2$. В результате операции расширения возникает структура, называемая нами цикл$_2$. Примеры на цикл$_2$ даны выше (см. стр. 33, 2, 3, 6).

Трансференция и цикл$_2$ суть однотактные структуры.

Цикл$_1$. В результате применения двухтактной операции (расширение-ЭС)$_{T_1}$ 2 (сжатие-ЧЭ)$_{T_2}$ возникает двухтактная структура, называемая нами цикл$_1$, например:

---

[7] Термины: «трансференция», «цикл$_1$», «цикл$_2$» употребляются в значении, близком, но не тождественном тому, в котором они употреблялись нами ранее (см. ВЯ, 1961, № 2).

$$\frac{\begin{array}{l}(a\,//\,b)\cdot(c\,//\,d)\\(a\,//\,b)\cdot(c\,//\,d)\\a=(a\,//\,b)\cdot(c\,//\,d)\end{array}}{\begin{array}{l}(((a\,//\,b)\cdot(c\,//\,d))\,//\,b)\cdot(c\,//\,d)\\(a\,//\,b)\cdot(c\,//\,d)\\a\,//\,b=a\\c\,//\,d=b\end{array}}\cdot$$
$$((a\cdot b)\,//\,b)\cdot(c\,//\,d)$$

Итерация. В результате двухтактной операции $(\text{расширение-ЭС})_{T_1}$ 2 $(\text{сжатие-ЧЭ})_{T_2}$ мы получаем двухтактную структуру, называемую нами итерацией в отличие от цикла$_1$; при подстановке 2 ЧЭ_{T_2} подставляемые элементы тождественны, например:

$$\frac{\begin{array}{l}(a\,//\,b)\cdot(c\,//\,d)\\(a\,//\,b)\cdot(c\,//\,d)\\a=(a\,//\,b)\cdot(c\,//\,d)\end{array}}{\begin{array}{l}(((a\,//\,b)\cdot(c\,//\,d))\,//\,b)\cdot(c\,//\,d)\\(a\,//\,b)\cdot(c\,//\,d)\\(a\,//\,b)=a\\(c\,//\,d)=a\end{array}}\cdot$$
$$((a\cdot a)\,//\,b)\cdot a$$

Иерархия РНДС. Структуру, полученную в результате n-тактной подстановки (∀ Э) $(\text{расширение-ЭС})_{T_k}$, $k=1, 2, 3, \ldots$, мы назовем иерархией РНДС, например:

основная структура $(a\,//\,b)\cdot(c\,//\,d)$;
исходная структура
$$\begin{array}{l}(a_1\,//\,b_1)\cdot(c_1\,//\,d_1)\\(a_2\,//\,b_2)\cdot(c_2\,//\,d_2)\\(a_3\,//\,b_3)\cdot(c_3\,//\,d_3)\\(a_4\,//\,b_4)\cdot(c_4\,//\,d_4)\\a=(a_1\,//\,b_1)\cdot(c_1\,//\,d_1)\\b=(a_2\,//\,b_2)\cdot(c_2\,//\,d_2)\\c=(a_3\,//\,b_3)\cdot(c_3\,//\,d_3)\\d=(a_4\,//\,b_4)\cdot(c_4\,//\,d_4)\end{array}$$

$$[[(a_1\,//\,b_1)\cdot(c_1\,//\,d_1)]\,//\,[(a_2\,//\,b_2)\cdot(c_2\,//\,d_2)]]\cdot([[(a_3\,//\,b_3)\,(c_3\,//\,d_3)]\,//\,[(a_4\,//\,b_4)\cdot(c_4\,//\,d_4)])$$

Полученная структура есть двухтактная иерархия.

Транспланации. Класс операций подстановок, выполняемых в структурах различных тактов одной иерархии РНДС, дает структуры, называемые нами транспланациями. Можно говорить о транспланационных трансференциях, транспланационных циклах и транспланационных итерациях.

Дадим сводную таблицу, определяющую типы ПДС через операцию подстановки (см. табл. 3).

Нам следует остановиться на представлении выделенных нами типов результирующих ПДС в терминах графов.

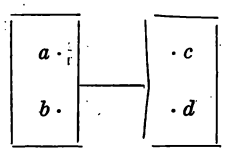
Трансференция, которая возникает в результате обмена, заменяет вершины, не затрагивая общего строения графа, непример:

$$\frac{\begin{array}{l}(a\,//\,b)\cdot(c\,//\,d)\\a=a_1\quad c=c_1\end{array}}{\begin{array}{l}b=b_1\quad d=d_1\\(a_1\,//\,b_1)\cdot(c_1\,//\,d_1)\end{array}}$$

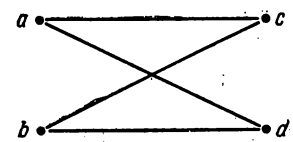
Таблица 3

Типы ПДС	Операции подстановки
Трансференция	(Обмен-ЭЭ, -ЧЧ, -СС)$_{T_1}$
Цикл$_2$	(Расширение-ЭЧ, -ЭС, -ЧС)$_{T_1}$
Цикл$_1$	(Расширение-ЭС)$_{T_1}N$ (сжатие-ЧЭ)$_{T_2}$ $Э_1 = Э_2 = \ldots, = Э_N$
Итерация	(Расширение-ЭС)$_{T_1}N$ (сжатие-ЧЭ)$_{T_2}$ $Э_1 = Э_2 = \ldots = Э_N$
n-тактная иерархия	(ЭЭ) (расширение-ЭС)$_{T_k}$ $k = 1, 2, 3, \ldots n$
Транспланационная трансференция	(ЭЭ) (расширение-ЭС)$_{T_k}$ (обмен-ЭЭ, -ЧЧ, -СС)$_{T_{k+1}}$
Транспланационный цикл$_2$	(ЭЭ) (расширение-ЭС)$_{T_k}$ (расширение-ЭЧ, -ЭС, -ЧС)$_{T_{k+1}}$
Транспланационный цикл$_1$	(ЭЭ) (расширение-ЭС)$_{T_{k+1}}N$ (сжатие-ЧЭ)$_{T_{k+2}}$ или (расширение-ЭС)$_{T_{k+1}}$ $(Э_1 \neq Э_2 \neq \ldots \neq Э_N)$ N (сжатие-ЧЭ)$_{T_{k+2}}$
Транспланационная итерация	(ЭЭ) (расширение-ЭС)$_{T_k}N$ (сжатие-ЧЭ)$_{T_{k+1}}$ $Э_1 = Э_2 = \ldots Э_N$ или (расширение-ЭС)$_{T_{k+1}}N$ (сжатие-ЧЭ)$_{T_{k+2}}$

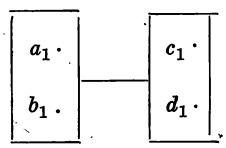
превращает граф



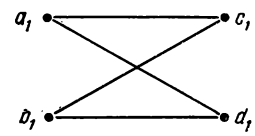
или



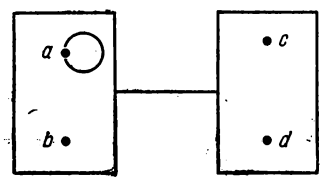
в граф



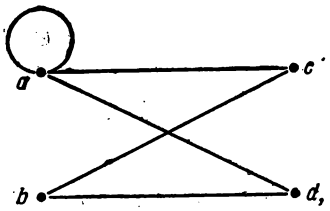
или



Цикл$_2$ рассматривался нами в примерах на расширение (см. стр. 34). Итерация, т. е. структура вида $((a \cdot a) // b) \cdot (c // d) = (a \cdot a \cdot c) // (b \times \times c) // (a \cdot a \cdot d) // (b \cdot d)$
дает следующий граф:

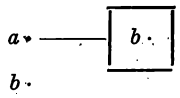


или



т. е. граф с одной (в данном случае) петлей (boucle). Цикл$_1$ вызывает большие трудности в записи.

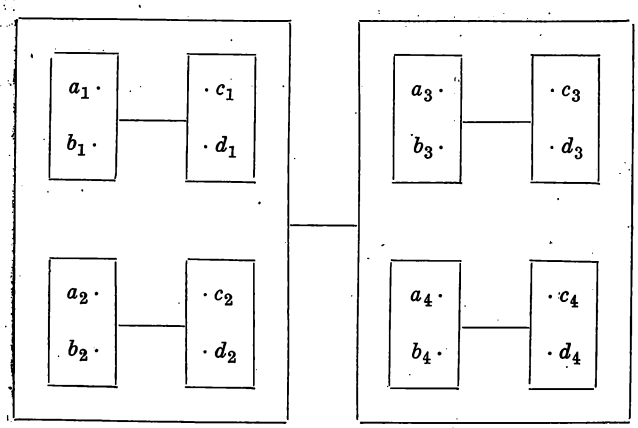
Действительно, цикл$_1$ характеризуется такой частью формулы: $(a \cdot b) // b$. В записи элемент b должен быть одновременно связан с a и отделен от a. Мы введем следующего рода запись



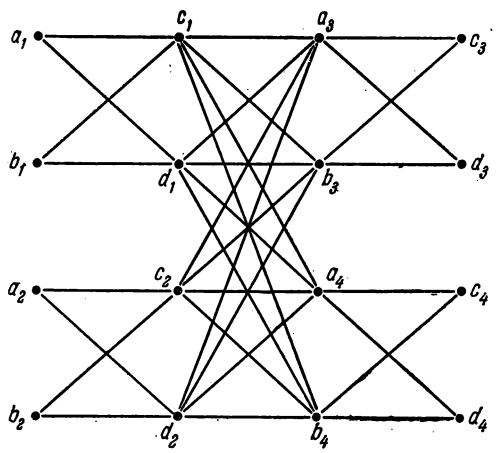
означающую, что b входит в отношение R_0 в результате вторичного отношения.

Иерархия РНДС в терминах графов примет следующий вид:

$$[((a_1 // b_1) \cdot (c_1 // d_1))_a // ((a_2 // b_2) \cdot (c_2 // d_2))_b] \cdot [((a_3 // b_3) \cdot (c_3 // d_3))_c // ((a_4 // b_4) \cdot (c_4 // d_4))_d]$$



или



Теперь мы должны остановиться на связи понятий подстановки и дистрибутивности. Можно ли получить из РНДС ее коррелят только посредством ряда подстановок?

Покажем, что справедливо следующее:

никакая подстановка или ряд подстановок не имеют результирующей коррелята РНДС и ОДС вообще. Рассмотрим всевозможные случаи:

1. Подстановка есть подстановка, заданная таблицей (обмен, расширение, сжатие).

а) Обмен. Легко увидеть, что обмен не нарушает распределения отношений R_0 и R_2 по скобкам:

(//)·(//).

б) Расширение (ЭЧ, ЭС, ЧС). Так как подстановка совершается на место элемента или члена РНДС, то отношение R_0 не затрагивается расширением:

( )·( ).

в) Сжатие (ЧЭ, СЭ, СЧ). Сжатие ЧЭ не по всем членам сохраняет отношение R_0. Сжатие ЧЭ по всем членам, сжатие СЭ и СЧ могут быть применены лишь к части РНДС (по определению ДС). Пусть в РНДС $(a\,//\,b)\cdot(c\,//\,d)\;(c^1\,//\,d^1)\;(a\,//\,b)\cdot(c\,//\,d)$ — часть, к которой применимо сжатие, $(c^1\,//\,d^1)$ — остаток; тогда при любом виде сжатия сохраняется отношение R_0, связывающее основную часть РНДС и остаток:

$(a\cdot b)\cdot(c^1\,//\,d^1)$ или $(a\,//\,b)\cdot(c^1\,//\,d^1)$ или $(a)\cdot(c^1\,//\,d^1)$.

2. Возьмем подстановки, образованные n-тактным применением простых подстановок. Можно сразу отбросить подстановку обмена как не меняющую структуру.

Остается рассмотреть две комбинации операций — $(\text{расширение})_{T_k}\;(\text{сжатие})_{T_{k+1}}$ и $(\text{сжатие})_{T_k}\;(\text{расширение})_{T_{k+1}}$:

а) $(\text{расширение})_{T_k}\;(\text{сжатие})_{T_{k+1}}$.

Как было показано выше, расширение не затрагивает отношений R в РНДС, а сжатие в силу определения ДС должно сохранить хотя бы одно из них.

б) $(\text{сжатие})_{T_k}\;(\text{расширение})_{T_{k+1}}$.

Сжатие сохраняет хотя бы одно отношение R_0, а расширение его не затрагивает.

Таким образом, невозможность появления ОДС (в частности, коррелята) в результирующей структуре доказана.

5.21. Неравномерно нормальная дистрибутивная структура (ННДС). Практически важным случаем ПДС является структура, отличающаяся от РНДС тем, что она допускает различное число термов по R_2 в сложных термах по R_0, например: $(a\,//\,b)\cdot(c\,//\,d\,//\,e)$.

Выше такая структура была получена нами путем расширения-ЭЧ (см. пример на стр. 33).

Назовем ПДС такого вида неравномерно нормальной дистрибутивной структурой (ННДС). Структуру, получаемую из ННДС в силу дистрибутивности R_0 относительно R_2, назовем корреляторм ННДС. Как и в случае с РНДС, укажем ряд соотношений, связывающих ННДС и ее корреляты.

Набор символов тот же, что и для РНДС: l_k — B-длина k-ой скобки ННДС.

Соотношение 1

$$l_1 = m$$

Соотношение 2

$$m_1 = \prod_{k=1,\ldots,m} l_k$$

Соотношение 3

$$\lambda = \sum_{k=1,\ldots,m;} l_k$$

частный случай:

$$\lambda = (l_{1-k}) \cdot k + l_{(k+1)}^{m}$$

Соотношение 4

$$\lambda_1 = m \cdot \prod l_k$$

Соотношение 5

$$p = \sum (l_k - 1)$$

Соотношение 6

$$r = m - 1$$

Соотношение 7

$$p_1 = \prod (l_k -$$

Соотношение 8

$$r_1 = \prod (l_k - 1)$$

Определение

$$L = \lambda + p + r + m = \sum l_k + \sum (l_k - 1) + 2m - 1$$

Определение

$$L_1 = \lambda_1 + p_1 + r_1 + m_1 = \prod l_k \cdot (2m + 1) - 1.$$

Определение

$$Э = \frac{L_1}{L} = \frac{\prod l_k \cdot (2m+1) - 1}{\sum l_k + \sum (l_k - 1) + 2m - 1}.$$

5.3. Обратные дистрибутивные структуры (ОДС)

Обратные дистрибутивные структуры возникают благодаря применению операции подстановки к равномощным кортежам множества J (т. е. к коррелятам РНДС).

В качестве примера мы возьмем наипростейший коррелят:

$$(a // b) \cdot (c // d) = (a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d).$$

Как и в случае РНДС, назовем терм по отношению $R_2 — (a \cdot c)$ элементом, а терм по отношению $R_0 — a$, c — членом (здесь и далее слово «элемент» употребляется в ином смысле, чем в начале работы: элемент множества).

Имеем таблицу, дающую типы подстановок (табл. 4).

Таблица 4

Первый тип \ Второй тип	Элемент	Член	Структура
Элемент	1. Обмен-ЭЭ	2. Сжатие-ЭЧ	3. Расширение-ЭС
Член	4. Расширение-ЧЭ	5. Обмен-ЧЧ	6. Расширение-ЧС
Структура	7. Сжатие-СЭ	8. Сжатие-СЧ	9. Обмен-СС

Примеры:

2. Сжатие-ЭЧ

а) $(a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d)$ (автоморф.)[8],

$$\frac{a \cdot c = a}{(a) // (b \cdot c) // (a \cdot d) // (b \cdot d)};$$

с извлечением за скобку общего терма по R_0 имеем:

$$(a \cdot d) // (b \cdot (c // d)).$$

б) $(a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d)$ (автоморф.),

$$\frac{a \cdot c = a;\ b \cdot c = b;\ a \cdot d = c;\ b \cdot d = d}{a // b // c // d}[9];$$

3. Расширение-ЭС

$(a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d)$ (гетероморф.),

$$\frac{a \cdot c = (a_1 \cdot c_1) // (b_1 \cdot c_1) // (a_1 \cdot d_1) // (b_1 \cdot d_1)}{(a_1 \cdot c_1) // (b_1 \cdot c_1) // (a_1 \cdot d_1) // (b_1 \cdot d_1) // (b \cdot c) // (a \cdot d) // (b \cdot d)};$$

с извлечением за скобку общего терма по R_0 имеем

$$[(a // b) \cdot (c // d)] // [(d \cdot (a // b) // (b \cdot c)].$$

4. Расширение-ЧЭ

а) $(a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d)$ (гетероморф.),

$$\frac{a = g \cdot h}{(g \cdot h \cdot c) // (b \cdot c) // (g \cdot h \cdot d) // (b \cdot d)};$$

вынося за скобки общий терм по R_0, имеем:

$$[(g \cdot h) // b] \cdot [c // d].$$

б) $(a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d)$ (гетероморф.),

$$\frac{a = g \cdot h;\ b = e \cdot f;\ c = i \cdot j;\ d = k \cdot l}{(g \cdot h \cdot i \cdot j) // (e \cdot f \cdot i \cdot j) // (g \cdot h \cdot k \cdot l) // (e \cdot f \cdot k \cdot l)};$$

вынося за скобки общий терм по R_0, имеем:

$(g \cdot h) // (e \cdot f) \cdot (i \cdot j) // (k \cdot l)$.

6. Расширение-ЧС

$(a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d)$ (гетероморф.),

$$\frac{a = (a_1 \cdot c_1) // (b_1 \cdot c_1) // (a_1 \cdot d_1) // (b_1 \cdot d_1)}{([(a_1 \cdot c_1) // (b_1 \cdot c_1) // (a_1 \cdot d_1) // (b_1 \cdot d_1)] \cdot c) // (b \cdot c) // ([(a_1 \cdot d_1) // (b_1 \cdot d_1)] \cdot d) // (b \cdot d)}.$$

Вынося за скобку общий терм по R_0, имеем:

$$\{[(a_1 // b_1) \cdot (c_1 // d_1)] // b\} \cdot \{c // d\}.$$

---

[8] В целях сокращения записи мы обозначаем исходную структуру, идентичную с основной, «автоморф.», а исходную структуру, отличную от основной, — «гетероморф.».

[9] В случае 2б результирующая структура может рассматриваться как ДС. Действительно, $(a // b // c // d) \equiv (a \cdot 0 // b \cdot 0 // c \cdot 0 // d \cdot 0)$ по определению нуля.

7. Сжатие-СЭ

$(a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d)$ (автоморф.),

$$\frac{(a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d) = a \cdot c}{a \cdot c}$$

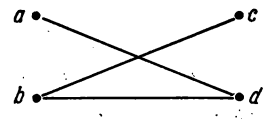
8. Сжатие-СЧ

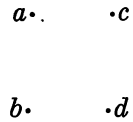
$(a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d)$ (автоморф.),

$$\frac{(a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d) = a}{a}$$ [10].

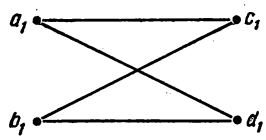
В форме графов полученные в результате подстановки структуры примут следующий вид:

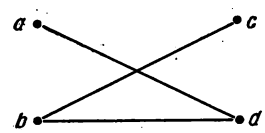
2. Сжатие-ЭЧ

(а) 

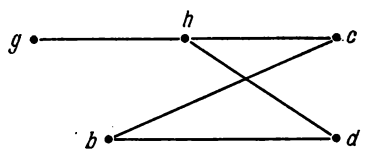
(б) 

3. Расширение-ЭС

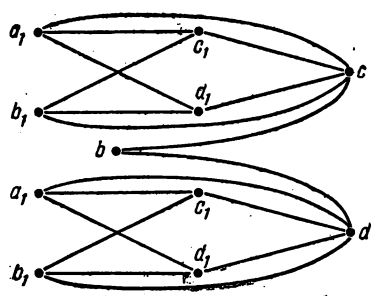




4. Расширение-ЧЭ



6. Расширение-ЧС



Сжатие-СЭ

[10] В случаях 7, 8 результирующая не является ДС и подстановки вида 7, 8 действительны лишь для основных структур, являющихся частью более сложной структуры.

Сжатие-СЧ

a

К ОДС применимы понятия, связанные с многократными подстановками.

Рассмотрим некоторые ОДС, получаемые путем многократных подстановок.

I. ОДС представляют собой род равномощных кортежей, не имеющих общих членов по R_0, например $(a'' \cdot b'' \cdot c'') // (d'' \cdot e'' \cdot f'')$.

Данная структура может быть получена из коррелята РНДС путем многократных подстановок, например:

$(a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d)$ (гетероморф.), (сжатие-ЭЧ)

$$\frac{a \cdot c = a^1 \quad a \cdot d = c^1 \qquad b \cdot c = b^1 \quad b \cdot d = d^1}{(a^1 // b^1) // (c^1 // d^1)}$$ (гетероморф.), (расширение $\underset{\nu>\mu}{\mu Э_\nu Ч}$)

$$\frac{a^1 // b^1 = a'' \cdot b'' \cdot c'' \qquad c^1 // d^1 = d'' \cdot e'' \cdot f'' \cdot}{(a'' \cdot b'' \cdot c'') // (d'' \cdot e'' \cdot f'')}$$

или $(a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d)$ (гетероморф.), (сжатие-ЭЧ)

$$\frac{a \cdot c = a^1 \quad a \cdot d = c^1 \qquad b \cdot c = b^1 \quad b \cdot d = d^1}{a^1 // b^1 // c^1 // d^1}$$ (гетероморф.), (расширение-ЧЭ)

$$\frac{a^1 = a'' \cdot b'' \quad c^1 = e'' \cdot f'' \qquad b^1 = c'' \cdot d'' \quad d^1 = g'' \cdot h''}{(a'' \cdot b'') // (c'' \cdot d'') // (e'' \cdot f'') // (g'' \cdot h'')}.$$

II. ОДС представляет собой ряд тождественных кортежей:
$(a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d)$ (автоморф.),

$$\frac{a = b; \quad c = d}{(a \cdot c) // (a \cdot c) // (a \cdot c) // (a \cdot c)}.$$

III. ОДС представляет собой ряд кортежей, имеющих N общих членов (пусть $Л$ — длина кортежа; $N < Л$):
$(a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d)$ (гетероморф.),

$$\frac{a \cdot c = a^1; \quad b \cdot c = b^1; \quad a \cdot d = c^1; \quad b \cdot d = d^1}{a^1 // b^1 // c^1 // d^1}$$ (гетероморф.),

$$\frac{a^1 = a'' \cdot b''; \quad c^1 = e'' \cdot f'' \qquad b^1 = c'' \cdot d''; \quad d^1 = g'' \cdot h''}{(a'' \cdot b'') // (c'' \cdot d'') // (e'' \cdot f'') // (g'' \cdot h'')}$$ (автоморф.),

$$\frac{b'' = d'' = f'' = h''}{(a'' \cdot b'') // (c'' \cdot b'') // (e'' \cdot b'') // (g'' \cdot b'')}.$$

при вынесении общего терма по R_0 за скобку имеем:
$b'' \cdot (a'' // c'' // e'' // g'')$.

Следует различать два случая:

а) $Л - N = 1$ (как в нашем примере), в этом случае имеем прямую ДС особого вида;

б) $Л - N > 1$, в этом случае имеем ОДС с вынесенными за скобку общими термами по R_0.

IV. ОДС представляет собой M_1 кортежей, из них k кортежей имеют h одних общих термов по R_0, а $M_1 - k$ имеют g других общих термов по R_0:

$(a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d)$ (гетероморф.),

$$\frac{\begin{array}{c} a \cdot c = a^1;\ a \cdot d = c^1 \\ b \cdot c = b^1;\ b \cdot d = d^1 \end{array}}{a^1 // b^1 // c^1 // d^1} \text{ (гетероморф.)},$$

$$\frac{\begin{array}{c} a^1 = a'' \cdot b'';\ c^1 = e'' \cdot f'' \\ b^1 = a'' \cdot d'';\ d^1 = g'' \cdot f'' \end{array}}{(a'' \cdot b'') // (a'' \cdot d'') // (e'' \cdot f'') // (g'' \cdot f'').}$$

По вынесении за скобку общих термов по R_0 имеем:
$a'' \cdot (b'' // d'') // f'' \cdot (e'' // g'')$.

Таким образом, в случае IV кортежи делятся на две группы (из k и $M_1 - k$ кортежей), связанных R_2. Каждая группа кортежей представляет случай III (см. выше).

V. ОДС представляет собой M_1 кортежей, из которых k кортежей включают h общих термов по R_0, а $M_1 - k$ кортежей включают g общих термов по R_0, причем некоторые термы по R_0 принадлежат и h и g.

$(a \cdot c) // (b \cdot c) // (a \cdot d) // (b \cdot d)$ (гетероморф.),

$$\frac{\begin{array}{c} a \cdot c = a_1 \cdot b_1 \cdot c_1 \\ b \cdot c = a_1 \cdot e_1 \cdot f_1 \\ a \cdot d = g_1 \cdot e_1 \cdot f_1 \\ b \cdot d = j_1 \cdot e_1 \cdot f_1 \end{array}}{(a_1 \cdot b_1 \cdot c_1) // (a_1 \cdot e_1 \cdot f_1) // (g_1 \cdot e_1 \cdot f_1) // (j_1 \cdot e_1 \cdot f_1)};$$

после вынесения общих термов по R_0 имеем:

$$a_1 \cdot ((b_1 \cdot c_1) // (e_1 \cdot f_1)) // (e_1 \cdot f_1 \cdot (g_1 // j_1))$$

или

$a_1 \cdot b_1 \cdot c_1 // e_1 \cdot f_1 \cdot (a_1 // g_1 // j_1)$, т. е. в случае V мы встречаемся с множественностью решений.

6. НЕДОСТАТОЧНОСТЬ И СИСТЕМНОСТЬ ДС

Для того, чтобы получить некоторые количественные характеристики ДС, введем два новых понятия.

Предварительно напомним о некоторых символах, введенных выше, и добавим к ним некоторые дополнительные символы:

L — длина РНДС,
L' — длина ННДС,
L'' — длина исследуемой ПДС или преобразованной ОДС,
L_1 — длина коррелята РНДС,
L'_1 — длина коррелята ННДС,
L''_1 — длина исследуемой ОДС.

Недостаточностью исследуемой ДС (U) назовем разность длин ДС и коррелята ННДС (РНДС), который может быть «надстроен» над данной ДС:

$$U = L'_1 - L''_1.$$

Получить ННДС (РНДС) и затем ее коррелят мы можем, объединив все элементы исследуемой ДС с одним и тем же индексом в один класс. Например, пусть имеется ДС:

$$(a_1 \cdot b_2) \,//\, (c_1 \cdot d_2).$$

«Надстроенной» РНДС будет:

$$(a_1 \,//\, c_1) \cdot (b_2 \,//\, d_2).$$

Ее коррелят будет:

$$(a_1 \cdot b_2) \,//\, (c_1 \cdot b_2) \,//\, (a_1 \cdot d_2) \,//\, (c_1 \cdot d_2)$$
$$U = L_1' - L_1''$$
$$L_1' = 8 + 4 + 4 + 3 = 19,$$
$$L_1' = 4 + 2 + 2 + 1 = 9,$$
$$U = 19 - 9 = 10.$$

Системностью (S) исследуемой ДС назовем отношение длин ННДС (или РНДС) и исследуемой ПДС (преобразованной ОДС):

$$S = \frac{L'}{L''}.$$

Под преобразованной ОДС мы понимаем ОДС, в которой некоторые общие термы по R_0 вынесены за скобку. Например, дана ДС:

$$(a_1 \cdot b_2 \cdot c_3 \cdot d_4) \,//\, (a_1 \cdot b_2' \cdot c_3 \cdot d_4') \,//\, (a_1 \cdot b_2'' \cdot c_3 \cdot d_4'') =$$
$$= (a_1) \cdot (c_3) \cdot ((b_2 \cdot d_4) \,//\, (b_2' \cdot d_4') \,//\, (b_2'' \cdot d_4'')),$$

соответствующая ННДС будет: $(a_1) \cdot (b_2 \,//\, b_2' \,//\, b_2'') \cdot (c_3) \cdot (d_4 \,//\, d_4' \,//\, d_4'')$

$$L' = 8 + 4 + 3 + 4 = 19,$$
$$L'' = 8 + 6 + 5 + 2 = 21,$$
$$S = \frac{L'}{L''} = \frac{19}{21} \approx 0{,}9.$$

Если исследуемая ДС является ННДС или РНДС, то ее $S = 1$, а $U = 0$.

ELEMENTS OF A THEORY OF THE LANGUAGE UNITS DISTRIBUTION

Summary

This paper deals with some notions of a meta-language for describing language structure.

Two relations between elements of a set of indexed sequences J are established: R_0 and R_2.

R_0 holds between elements (or groups of elements) belonging to the same sequence and having indexes k and $k+1$ (i. e. differing from one another by one).

R_2 holds between elements (or groups of elements) belonging to different sequences but having the same index.

The relations R_0 and R_2 correspond to the syntagmatic and paradigmatic relations of post-Saussurian linguistics.

Both relations are tied up by the distributional law:

$$aR_0(bR_2c) = (aR_0b)\,R_2\,(aR_0c) \qquad (1)$$

Any sequense of elements including the both relations may be called a distributive structure. The distributive structure represented by the

left part of (1) is called a direct distributive structure. The distributive structure represented by the right part of (1) is called a conversed distributive structure. The distributive structure (direct or conversed) any relationterm of which includes an equal number of elements is called a regular distributive structure.

A broad class of substitutions is proposed to generate various kinds of distributive structures from a regular distributive structure.

The author considers the ideas expounded in the paper to be his first approach to establishing a general calculus for language structure.

С. Маркус

ЛОГИЧЕСКИЙ АСПЕКТ ЛИНГВИСТИЧЕСКИХ ОППОЗИЦИЙ

I. ТЕОРЕТИКО-МНОЖЕСТВЕННАЯ ОСНОВА ЛИНГВИСТИЧЕСКИХ ОППОЗИЦИЙ

Первым лингвистом, обратившим внимание на центральную роль, которую играет понятие оппозиции в исследовании языка, был Фердинанд де Соссюр [12]. Поставленный Ф. де Соссюром вопрос о систематическом изучении лингвистических оппозиций был разработан Н. С. Трубецким, который осуществил классификацию различных типов оппозиций в фонологии, указав одновременно на возможность применения основных принципов его классификации и к другим областям лингвистики [13]. Классификация Н. С. Трубецкого была улучшена Ж. Кантино. Он систематически исследовал возможность применения классификации Н. С. Трубецкого к так называемым «сигнификативным оппозициям» [26] и показал, что различные типы оппозиций, обнаруженные Н. С. Трубецким, соответствуют фундаментальным отношениям в символической логике [27]. Ж. Кантино предложил использовать термин «отношение» вместо «оппозиция», так как первый термин, будучи более общим, адекватнее. Однако, поскольку термин «оппозиция» широко употребляется в литературе последних лет, мы будем пользоваться обоими терминами, считая их эквивалентными (так поступает и Ж. Кантино). Изучение лингвистических отношений продвинулось вперед благодаря работам А. Мартине [40], Л. Ельмслева [4], Л. Теньера [46], а также американских дескриптивистов (например, З. Харриса [34]) и т. д. Была проведена интересная аналогия между типами лингвистических оппозиций и некоторыми ситуациями в теории кодов [23].

В исследованиях последних лет по структурному синтаксису часто упоминаются отношения упорядоченности и квазиупорядоченности в смысле [24] (см., например, [17], [46], [34] и [6]), а при выяснении сущности понятия фонемы как понятия, принадлежащего высшему уровню познания, была констатирована полезность использования аппарата теории пропозиционных функций с двумя переменными (см. С. К. Шаумян: [18], [19], [20]).

За последние годы занятия кибернетической лингвистикой вызвали интерес к формальному анализу языка. В то же время формальный анализ применяют ко все более сложным лингвистическим отношениям. Это требует усовершенствования метода исследования, более тесной его связи с логико-математическими аспектами лингвистических отношений. Необходимо заново проанализировать эти отношения, используя при этом не только отношения включения и элементарные операции теории множеств, как представляет себе Ж. Кантино, но и другие типы отношений и операции математики, все разнообразие структур, изучаемых в современной матаматике. Это изучение потребует применения

и обычного в математике аппарата доказательства, чтобы можно было установить логические следствия лингвистических отношений, свойства этих отношений.

Можно также попытаться выполнить часть указанной выше программы, а именно ту часть, которая касается оппозиций, описываемых посредством отношения включения, элементарных операций теории множеств и отношений эквивалентности.

Далее, как показано в [8], описание определенных типов лингвистических отношений требует использования более общих бинарных отношений. Этот вопрос мы рассмотрим в другой статье.

1. Некоторые понятия теории множеств

Слово «множество» сохраняет в нашей статье свой обычный смысл: оно обозначает первичное понятие, не сводимое к более простым понятиям.

Предметы, составляющие множество, называются элементами множества.

Отношение «a есть элемент множества A» называется отношением принадлежности и обозначается знаком $\in$ (знак Дж. Пеано): $a \in A$. Если элемент b не принадлежит множеству A, пишем $b \notin A$.

Как определяются множества? Существуют два пути:

а) наименование всех его элементов (перечисление), например: множество, образованное числами 2, 5, 9, 11, 13, 14; множество, образованное словами *человек*, *стекло*, *тетрадь*;

б) указание характерных свойств элементов (описанием), например: множество имен существительных русского языка, множество правильно построенных предложений румынского языка, множество четных чисел, множество падежей немецкого языка.

Множество, определенное перечислением элементов, может состоять только из конечного числа элементов — именно потому, что оно образовано их перечислением. Множество, определенное указанием свойств, может быть как конечным, так и бесконечным. Например, множество падежей румынского языка конечно; множество действительных чисел бесконечно.

Итак, в отличие от бесконечных множеств, которые определяются только описанием, конечные множества определяются как перечислением элементов, так и описанием. Тот факт, что некоторые конечные множества описываются одним из этих двух процессов, зависит от состояния наших знаний.

Некоторые множества, определенные посредством описания свойств, могут быть определены и посредством перечисления.

Рассмотрим, например, следующее множество морфем: {*a*, *ă*, *e*, *ei*, *ele*, *elor*}. Это множество определяется перечислением элементов. Но если мы свяжем это множество с окончаниями румынского слова *casă* 'дом', то станет ясно, что это множество можно определить и таким образом: множество суффиксальных морфем, которые выступают в качестве окончаний слова *casă*.

Определение множества посредством указания свойств элементов выше, с научной точки зрения, нежели определение посредством перечисления, так как оно не ограничивается тем, что очевидно на первый взгляд, а вскрывает нечто из сущности, общей для элементов множества.

Определение множества указанием признаков его элементов может нередко осуществляться несколькими различными способами, так как очень часто для элементов некоторого множества оказываются общими одновременно несколько различных свойств. Например, множество {1, 2, 3, 4, 5} — это, с одной стороны, множество натуральных чисел, мень-

ших числа 6, а с другой стороны, это множество, получаемое вычитанием числа 6 из натуральных чисел, бо́льших 6 и меньших 12.

Точно так же множество морфем {*a*, *ă*, *e*, *ei*, *ele*, *elor*} — это вместе с тем и множество суффиксальных морфем, появляющихся в качестве окончаний у румынского слова *plasă* 'сетка'. Это приводит к мысли, что слова *plasă* и *casă* нужно сгруппировать друг с другом с точки зрения их окончаний. Иными словами, определение некоторых множеств посредством указания свойств их элементов дает возможность создания все более общих, все более абстрактных и все более простых группировок. Когда множество определяется через свойство элементов, следует помнить, что это свойство действует как фильтр, как сито. Для каждого объекта нужно ответить на вопрос, обладает ли этот объект соответствующим свойством. В зависимости от положительного или отрицательного ответа на этот вопрос объект принадлежит или не принадлежит множеству. Если рассматриваемое свойство такого рода, что для некоторых объектов не может быть определенного ответа, тогда это свойство неспособно определять множество. Например, если для некоторых слов мы не можем определить, являются ли они прилагательными или числительными, тогда мы не имеем права говорить о «множестве прилагательных». Точно так же, если о некоторых дифтонгах мы не можем сказать, образуют ли они одну или две фонемы, тогда мы не имеем права говорить о «множестве фонем». Эта констатация есть фундаментальное следствие, касающееся перспектив применения теории множеств в лингвистике. Это применение теории множеств плодотворно в той мере, в какой предварительно делаются упомянутые уточнения. В математике множество есть совокупность объектов, такая, что каждый объект либо входит, либо не входит в множество (и нет объектов, которые одновременно входили бы и не входили в множество).

Мы можем рассматривать множества, состоящие из одного элемента, например, множество, образованное единственным элементом *a*. Такое множество обозначаем через {*a*}. Мы можем также определить п у с т о е м н о ж е с т в о, т. е. множество, не содержащее ни одного элемента. Мы будем обозначать его через ø или 0.

Мы можем рассматривать множество, элементы которого в свою очередь являются множествами. Например, можно рассматривать множество всех шаров; но шар в свою очередь есть множество, а именно множество всех точек, находящихся на заданном расстоянии от заданной точки.

Другой пример. Мы можем рассматривать множество фонем румынского языка {*p*, *f*, *r*}; но каждая из этих фонем в свою очередь есть множество, а именно множество дифференциальных (или фонологических релевантных) признаков. Так, *p* = {смычность, глухость, двугубность}, *f* = {фрикативность, глухость, лабиодентальность}, *r* = {дрожащая артикуляция, нейтральность в отношении глухости—звонкости, средненёбность}.

2. Отношения между множествами

Пусть *X* — множество. Пусть все другие рассматриваемые здесь множества будут содержать только элементы множества *X*. Будем называть такое *X* н а д м н о ж е с т в о м.

Пусть *A* и *B* — множества. Предположим, что имеет место следующая ситуация: если $a \in A$, то $a \in B$.

В этом случае мы говорим, что множество *A* с о д е р ж и т с я в множестве *B* или в к л ю ч е н о в множество *B*, и отношения между *A* и *B* обозначаются термином «включение» и знаком $\subseteq$. Если

$$A \subseteq B, \tag{1}$$

мы говорим, что A есть подмножество или часть множества B.

ПРИМЕР. Пусть $X =$ {множество морфологических категорий}; $A =$ $=$ {число, падеж}, $B =$ {число, падеж, род, степень сравнения}. Имеет место отношение $A \subseteq B$, так как оба элемента множества A входят в множество B. Это отношение включения имеет лингвистическую содержательную интерпретацию, если мы имеем в виду, в частности, русский язык. Можно считать, что A представляет имя существительное, а B представляет имя прилагательное. Заметим, что множество B не включено в множество A, так как некоторые элементы B не принадлежат множеству A. Запишем это следующим образом:

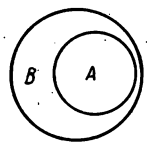
$$B \not\subseteq A. \tag{2}$$

Всякий раз, когда выполняются условия (1) и (2), мы будем говорить, что множество A строго включено или строго содержится в множестве B, и будем записывать это следующим образом:

$$A \subset B. \tag{3}$$

Отношение строгого включения A и B мы можем выразить и по-другому: отношение — или оппозиция — между A и B привативно не в пользу A (или привативно в пользу B). Например, оппозиция между именем существительным и именем прилагательным привативна в пользу имени прилагательного.

Мы можем наглядно представить отношение строгого включения между A и B следующим образом:



Отношение строгого включения обладает следующим свойством: Если оппозиция между A и B привативна в пользу B, а оппозиция между B и C привативна в пользу C, то оппозиция между A и C привативна в пользу C.

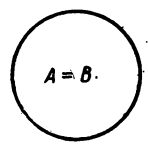
ДОКАЗАТЕЛЬСТВО. Допустим, что $A \subset B$ и $B \subset C$. Требуется доказать, что $A \subset C$. Пусть $a \in A$. Из отношения $A \subset B$ следует, что $a \in B$; из отношения $B \subset C$ следует, что $a \in C$, т. е. $A \subseteq C$. Но, с другой стороны, из $C \not\subseteq B$ следует, что существует элемент c такой, что $c \in C$, но $c \not\in B$. Из последнего отношения и из отношения $A \subset B$ следует, что $c \not\in A$, т. е. существует элемент, который принадлежит C, но не принадлежит A. Это значит $C \not\subseteq A$, откуда $A \subset C$. Доказательство закончено.

Пусть A и B — два множества, такие, что между ними имеет место отношение (1), но не имеет места отношение (3). В этом случае мы будем говорить, что множества A и B равны, и будем записывать это так:

$$A = B. \tag{4}$$

Отношение равенства между двумя множествами называется также нулевой оппозицией.

Она изображается так:



ПРИМЕР. Пусть A — множество шипящих согласных румынского языка, а B — множество таких согласных румынского языка, которые одновременно являются шипящими и средненёбными: $A = \{ş, j\}$, $B = \{ş, j\}$, т. е. $A = B$. Это объясняется тем фактом, что в румынском языке всякий шипящий согласный обязательно является средненёбным.

3. Операции с множествами

Будем называть объединением множеств A и B множество C, которое определяется следующим образом: элемент надмножества X принадлежит множеству C тогда и только тогда, когда этот элемент принадлежит по меньшей мере одному из множества A и B. Операция объединения обозначается знаком $\cup$. Поэтому

$$C = A \cup B.$$

Но определение объединения симметрично в отношении обоих его членов A и B, и поэтому

$$A \cup B = B \cup A.$$

Это есть свойство коммутативности объединения.

ПРИМЕР 1. $X = \{a, b, c, d, e, f\}$, $A = \{a, b, c, d, e\}$, $B = \{a, d, e, f\}$, $C = \{a\}$, $D = \{f\}$. $A \cup B = \{a, b, c, d, e, f\}$. $A \cup C = \{a, b, c, d, e\}$, $C \cup D = \{a, f\}$.

ПРИМЕР 2. X — множество имен существительных, A — множество имен существительных, не имеющих единственного числа, B — множество имен существительных, не имеющих множественного числа, $A \cup B$ — множество существительных, не имеющих полной парадигмы числа.

ПРИМЕР 3. X — множество согласных румынского языка, A — множество глухих согласных, B — множество смычных согласных, $A \cup B$ — множество глухих или смычных согласных: p, f, t, s, $ş$ k, m, b, n, d, g (p — глухой и смычный, f — глухой, но не смычный, g — смычный, но не глухой).

Будем называть пересечением или общей частью двух множеств A и B множество D, определенное следующим образом: элемент надмножества X принадлежит множеству D тогда и только тогда, когда этот элемент принадлежит как множеству A, так и множеству B. Операция пересечения обозначается через $\cap$. Следовательно,

$$D = A \cap B.$$

По определению пересечение множеств симметрично относительно обоих своих членов — A и B, и поэтому

$$A \cap B = B \cap A.$$

Это есть свойство коммутативности пересечения.

ПРИМЕР 1. $X = \{1, 2, 3, 4, 5, 6\}$; $A = \{1, 2, 3\}$; $B = \{2, 3, 4, 5\}$; $A \cap B = \{2, 3\}$.

ПРИМЕР 2. X — множество имен существительных и глаголов; A — множество слов надмножества X, изменяющихся по числам; B — множество слов надмножества X, изменяющихся по падежам; $A \cap B$ — множество имен существительных, которые не являются ни singularia tantum, ни pluralia tantum.

ПРИМЕР 3. X — множество морфологических категорий; A — множество морфологических категорий имени существительного; B — множество морфологических категорий глагола; $A \cap B = \{\text{число}\}$.

ЗАМЕЧАНИЕ. Операции объединения и пересечения распространяются на три и более множеств следующим образом. Если дано много

множеств: A_1, A_2, ..., A_n, ..., мы говорим, что множество A есть их объединение, и пишем

$$A = A_1 \cup A_2 \cup \ldots \cup A_n \cup \ldots,$$

если в A входят те и только те элементы, которые принадлежат по меньшей мере одному из множеств A_1, A_2, ..., A_n, ...

Мы говорим, что множество B есть пересечение или общая часть множеств A_1, A_2, ..., A_n, ..., и пишем

$$B = A_1 \cap A_2 \cap \ldots \cap A_n \cap \ldots,$$

если в B входят те и только те элементы, которые принадлежат одновременно всем множествам A_1, A_2, ..., A_n ...

Будем называть разностью двух множеств A и B множество E, определенное следующим образом: элементы надмножества X принадлежат множеству E тогда и только тогда, когда они принадлежат множеству A, но не принадлежат множеству B. Операция нахождения разности обозначается знаком —. Поэтому пишем:

$$E = A - B.$$

ПРИМЕР 1. $A = \{a, b, d, h\}$; $B = \{b, d\}$; $A - B = \{a, h\}$.

ПРИМЕР 2. A — множество осмысленных фраз, B — множество неосмысленных фраз, $A - B$ — это множество осмысленных фраз, т. е. A.

ПРИМЕР 3. A — множество следующих глаголов румынского языка: *a lucra* 'работать', *a mînca* 'есть', *a umbla* 'ходить'; B — множество глаголов, у которых по форме совпадают 3-е лицо единственного числа настоящего времени изъявительного наклонения и 3-е лицо множественного числа настоящего времени изъявительного наклонения; $A - B = \emptyset$.

Разность $X - B$ называется дополнением множества B. Будем обозначать его знаком $\bar{B}$.

ПРИМЕР 1. X — множество согласных; B — множество глухих согласных, $\bar{B}$ — множество звонких или нейтральных согласных.

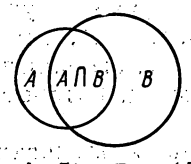
ПРИМЕР 2. X — множество слов, B — множество слов, которые не изменяются по временам, $\bar{B}$ — множество глаголов.

ПРИМЕР 3. X — множество ртовых гласных французского языка ($\{a, \varepsilon, e, i, ɔ, o, u, ø, œ, y\}$); $B = \{a, i, ɔ, o\}$; $\bar{B} = \{\varepsilon, e, u, ø, œ, y\}$.

4. Другие отношения между множествами

Пусть A и B — множества. Помимо рассмотренных выше отношений строгого включения и равенства, можно рассмотреть еще следующие типы отношений:

1. Если $A \neq \emptyset$, $B \neq \emptyset$, $A \not\subset B$, $B \not\subset A$ и $A \cap B \neq \emptyset$, тогда мы говорим, что множества A и B находятся в отношении эквиполентности, или в эквиполентной оппозиции. Это отношение наглядно изображается следующей схемой:



ПРИМЕР 1. $A = \{1, 2, 3, 4, 5\}$; $B = \{4, 5, 6\}$; $A \cap B = \{4, 5\}$; $B \not\subset A$, потому что $B - A = \{6\} \neq 0$, $A \not\subset B$, потому что $A - B = \{1, 2, 3\} \neq 0$.

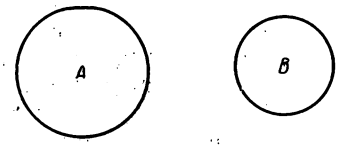
ПРИМЕР 2. X — множество морфологических категорий, $A = \{$число, род, падеж, степень сравнения$\}$, $B = \{$число, лицо, время, наклонение,

залог, вид}, $A \cap B = \{$число$\}$, $B \not\subset A$, так как $B - A = \{$лицо, время, наклонение, залог, вид$\}$, $A \not\subset B$, так как $A - B = \{$род, падеж, = степень сравнения$\} \neq \emptyset$. A и B находятся в эквиполентной оппозиции. В языке это оппозиция между именем прилагательным и глаголом.

ПРИМЕР 3. X — множество морфологических значений; $A = \{$единственное число, именительный падеж, определенный статус [1]$\}$, $B = \{$множественное число, именительный падеж, определенный статус$\}$, $A \cap B = \{$именительный падеж, определенный статус$\}$, $A - B = \{$единственное число$\}$, $B - A = \{$множественное число$\}$. Оппозиция между множествами A и B обобщает и синтезирует отношения между *casă* 'дом' и *casele* 'дома', между *pomul* 'дерево' и *pomii* 'деревья', между *caietul* 'тетрадь' и *caietele* 'тетради' и т. д. Это отношение эквиполентности.

2. Если $A \neq \emptyset$, $B \neq \emptyset$ и $A \cap B = \emptyset$, тогда мы говорим, что A и B находятся в отношении дизъюнкции и. Говорят также, что оппозиция между A и B дизъюнктивная. Некоторые авторы называют дизъюнктивную оппозицию отношением внешности (фр. exteriorité, рум. exterioritate): см., например, Ж. Кантино [27].

Это отношение наглядно изображается следующей схемой:



ПРИМЕР 1. $A = \{a, b, c\}$; $B = \{f, h\}$; $A \cap B = \emptyset$.

ПРИМЕР 2. A — множество губных согласных, B — множесто зубных согласных, $A \cap B = \emptyset$.

ПРИМЕР 3. $A = \{$единственное число, 1-е лицо, настоящее время, индикатив, действительный залог$\}$, $B = \{$множественное число, 3-е лицо, перфект, субъюнктив, страдательный залог$\}$, $A \cap B = \emptyset$. A и B — множество морфем в смысле Л. Ельмслева [38].

5. Анализ оппозиций

Любой тип отношений между двумя множествами A и B может быть охарактеризован посредством трех множеств: $A - B$, $B - A$ и $A \cap B$.

Оппозиция привативна тогда и только тогда, когда пусто одно и только одно из множеств $A - B$ и $B - A$. Оппозиция является нулевой тогда и только тогда, если $A - B = B - A = \emptyset$. Оппозиция эквиполентна тогда и только тогда, если $A \neq \emptyset$, $B \neq \emptyset$, $A \cap B \neq \emptyset$, $A - B \neq \emptyset$, $B - A \neq \emptyset$. Оппозиция дизъюнктивна тогда и только тогда, если $A \neq \emptyset$, $B \neq \emptyset$ и $A \cap B = \emptyset$. Получим следующую картину:

Таблица 1

Тип оппозиции	A	B	$A-B$	$B-A$	$A \cap B$
Привативная в пользу B			$= \emptyset$	$\neq \emptyset$	$= A$
Привативная в пользу A			$\neq \emptyset$	$= \emptyset$	$= B$
Нулевая			$= \emptyset$	$= \emptyset$	$= A = B$
Эквиполентная	$\neq \emptyset$	$\neq \emptyset$	$\neq \emptyset$	$\neq \emptyset$	$\neq \emptyset$
Дизъюнктивная	$\neq \emptyset$	$\neq \emptyset$	$= A$	$= B$	$= \emptyset$

[1] Определенный статус румынского языка (так называемый постпозитивный артикль), выражаемый особыми окончаниями имени, соответствует артиклю в западноевропейских языках; *casa* = *the house*, *pomul* = *the tree* (ср. *casă* = *house*, *pom* = *tree*).

Оппозиция между A и B, при которой $A \neq \varnothing \neq B$, называется собственной оппозицией; в другом случае она называется несобственной.

Легко видеть, что несобственная оппозиция — это либо нулевая, либо привативная оппозиция. Эквиполентная и дизъюнктивная оппозиции всегда собственны.

Множества A и B называются членами оппозиции.

Множества $A \cap B$ называются основанием сравнения множеств A и B.

$A - B$ и $B - A$ называются разностными множествами оппозиции между A и B. Мы можем поэтому говорить, что привативная оппозиция характеризуется тем, что одно и только одно из разностных множеств пусто. Эквиполентная оппозиция имеет место между двумя непустыми множествами и характеризуется тем, что как основание сравнения, так и разностные множества непусты. Дизъюнктивная оппозиция может существовать только между непустыми множествами и характеризуется тем, что основание сравнения пусто.

Введенные выше типы оппозиции восходят к классификации, которую дал лингвистическим оппозициям Н. С. Трубецкой [13].

Оппозиции интересны тогда, когда они имеют место между множествами, сходными в чем-либо между собой, т. е. между теми множествами, у которых есть общие элементы. Полное сходство делает оппозицию тривиальной (нулевая оппозиция); отсутствие частичного сходства, т. е. общих элементов, также уменьшает интерес к изучению оппозиции (дизъюнктивная оппозиция).

Таким образом, интереснее изучать оппозицию между множествами $A = \{$глухость, зубная артикуляция, фрикативность$\}$ и $B = \{$звонкость, зубная артикуляция, фрикативность$\}$, нежели оппозицию между $A = \{$глухость, зубная артикуляция, фрикативность$\}$ и $C = \{$звонкость, велярность, смычность$\}$, так как первая оппозиция имеет непустое основание сравнения, а вторая оппозиция имеет пустое основание. Первая оппозиция (оппозиция румынских фонем s и z) эквиполентна, вторая (оппозиция между s и g) дизъюнктивна.

6. Отношения между оппозициями

В дальнейшем будем обозначать оппозицию между двумя множествами A и B следующим образом: (A/B). Оппозиция между двумя множествами есть упорядоченная пара множеств. Следует различать оппозицию (A/B) и оппозицию (B/A).

Пусть дано множество $\mathfrak{A}$, элементы которого суть оппозиции. Если это оппозиции между множествами, являющимися элементами множества Ω, мы будем говорить, что $\mathfrak{A}$ есть множество оппозиций на Ω.

$\mathfrak{A}$ — это могут быть, например, все оппозиции, которые устанавливаются в румынском языке между множествами дифференциальных фонологических признаков (в этом случае будем обозначать такое $\mathfrak{A}$ через $\mathfrak{A}_{ф}$). $\mathfrak{A}$ могло бы быть также множеством оппозиций, которые в румынском языке устанавливаются между морфологическими дифференциальными признаками (в этом случае будем обозначать такое $\mathfrak{A}$ через $\mathfrak{A}_{м}$).

Пусть (A_1/B_1) и (A_2/B_2) — два элемента множества $\mathfrak{A}$. Мы будем говорить, что оппозиция (A_1B_1) пропорциональна (A_2/B_2) и будем писать $(A_1/B_1) \sim (A_2/B_2)$, если $A_1 - B_1 = A_2 - B_2$ и $B_1 - A_1 = B_2 - A_2$. Следовательно, две пропорциональные оппозиции имеют те же разностные множества.

ПРИМЕР 1. Элементы множества $\mathfrak{A}$ — это оппозиция между конечными множествами цветов. $A_1 = \{$красный, черный, желтый, синий$\}$,

B_1 = {синий, черный, фиолетовый, желтый}, A_2 = {красный, синий, белый, оранжевый}, B_2 = {белый, оранжевый, зеленый, фиолетовый}. Получаем $A_1 - B_1$ = {красный, синий} = $A_2 - B_2$, $B - A$ = {зеленый, фиолетовый} = $B_2 - A_2$. Следовательно, опозиция (A_1/B_1) пропорциональна оппозиции (A_2/B_2).

ПРИМЕР 2. $\mathfrak{A} = \mathfrak{A}_M$. A_1 = {единственное число, именительный падеж, определенный статус}, B_1 = {множественное число, именительный падеж, определенный статус}, A_2 = {единственное число, родительный падеж, определенный статус}, B_2 = {множественное число, родительный падеж, определенный статус}. Отсюда $A_1 - B_1$ = {единственное число} = $A_2 - B_2$. $B_1 - A_1$ = {множественное число} = $B_2 - A_2$. Следовательно, оппозиция (A_1/B_1) пропорциональна оппозиции (A_2/B_2).

ПРИМЕР 3. $\mathfrak{A} = \mathfrak{A}_\Phi$. A_1 = {фрикативность, глухость, зубная артикуляция}, B_1 = {фрикативность, звонкость, зубная артикуляция}, A_2 = {смычность, глухость, зубная артикуляция}, B_2 = {смычность, звонкость, зубная артикуляция}. Отсюда $A_1 - B_1$ = {глухость} = $= A_2 - B_2$; $B_1 - A_1$ = {звонкость} = $B_2 - A_2$. Следовательно, оппозиция (A_1/B_1) пропорциональна оппозиции (A_2/B_2). Мы можем поэтому говорить, что оппозиция между румынскими фонемами *s* и *z* пропорциональна оппозиции между *t* и *d*.

В связи с пониманием фонемы как множества дифференциальных признаков следует заметить, что понятие дифференциального признака берется здесь в реляционном смысле. См. по этому вопросу аргументацию С. К. Шаумяна в [18].

Оппозиция в множестве $\mathfrak{A}$, которая не пропорциональна никакой другой оппозиции в $\mathfrak{A}$, является, по определению, изолированной в $\mathfrak{A}$. Такая оппозиция представляет несравненно меньший научный интерес, нежели пропорциональная оппозиция.

ПРИМЕР. $\mathfrak{A} = \mathfrak{A}_\Phi$. A = {непрерывность, нейтральность в отношении глухости — звонкости, латериальность}, B = {фрикативность, нейтральность в отношении звонкости, передненёбность}. Оппозиция между A и B (т. е. оппозиция между согласным *l* и полусогласным *y* в румынском языке) изолирована в $\mathfrak{A}_\Phi$, так как в румынском языке не существует ни одного латерального согласного, кроме *l*, и ни одного передненёбного согласного, кроме *y*.

Мы будем говорить, что две оппозиции (A_1/B_1) и (A_2/B_2) пропорциональны влево, если $A_1 - B_1 = A_2 - B_2$.

ПРИМЕР 1. A_1 = {красный, желтый, зеленый}, B_1 = {красный, желтый, белый}, A_2 = {красный, желтый, зеленый}, B_2 = {красный, желтый, фиолетовый}. Отсюда $A_1 - B_2$ = {зеленый} = $A_2 - B_2$. Следовательно, оппозиция (A_1/B_1) пропорциональна влево оппозиции (A_2/B_2). Но (A_1/B_1) не пропорциональна (A_2/B_2), так как $B_1 - A_1$ = {белый} $\neq$ $\neq$ {фиолетовый} = $B_2 - A_2$.

ПРИМЕР 2. A_1 = {единственное число, именительный падеж, определенный статус}, B_1 = {единственное число, родительный падеж, определенный статус}, A_2 = {множественное число, именительный падеж, неопределенный статус}, B_2 = {множественное число, винительный падеж, неопределенный статус}. Отсюда $A_1 - B_1$ = {именительный падеж} = $A_2 - B_2$. Но (A_1/B_1) не пропорциональна (A_2/B_2), так как $B_1 - A_1$ = {родительный падеж} $\neq$ {винительный падеж} = $B_2 - A_2$, A_1, B_1, A_2 и B_2 — здесь множества морфем в смысле Л. Ельмслева [38].

Теорема 1. Если оппозиция (A_1/B_1) привативна в пользу A_1, а (A_2/B_2) привативна в пользу A_2 и если (A_1/B_1) пропорциональна влево оппозиции (A_2/B_2), в этом случае оппозиция (A_1/B_1) пропорциональна оппозиции (A_2/B_2).

ДОКАЗАТЕЛЬСТВО. По предположению о привативности в пользу первого члена, $B_1 - A_1 = B_2 - A_2 = \emptyset$. С другой стороны, по предположению о пропорциональности влево, $A_1 - B_1 = A_2 - B_2$. Следовательно, (A_1/B_1) пропорциональна оппозиции (A_2/B_2).

ТЕОРЕМА 2. Если (A_1/B_1) и (A_2/B_2) привативны не в пользу первого члена, тогда они пропорциональны влево.

ДОКАЗАТЕЛЬСТВО. Действительно, $A_1 - B_1 = A_2 - B_2 = \emptyset$.

ТЕОРЕМА 3. Если оппозиция (A_1/B_1) пропорциональна влево оппозиции (A_2/B_2) и если оппозиция (A_2/B_2) привативна не в пользу A_2, тогда (A_1/B_1) привативна не в пользу A_1.

ДОКАЗАТЕЛЬСТВО. По предположению, $A_1 - B_1 = A_2 - B_2$ и $A_2 - B_2 = \emptyset$ откуда $A_1 - B_1 = \emptyset$.

Аналогичным образом можно определить оппозицию, пропорциональную вправо. Оппозиция (A_1/B_1) пропорциональна вправо оппозиции (A_2/B_2), если $B_1 - A_1 = B_2 - A_2$.

ПРИМЕР. $A_1 = \{$единственное число, именительный падеж, определенный статус$\}$, $B_1 = \{$единственное число, родительный падеж, определенный статус$\}$, $A_2 = \{$множественное число, дательный падеж, неопределенный статус$\}$, $B_2 = \{$множественное число, родительный падеж, неопределенный статус$\}$. Отсюда $B_1 - A_1 = \{$родительный падеж$\} = B_2 - A_2$. Следовательно, оппозиция (A_1/B_1) пропорциональна вправо оппозиции (A_2/B_2). Однако оппозиция (A_1/B_1) не пропорциональна влево оппозиции (A_2/B_2), так как $A_1 - B_1 = \{$именительный падеж$\} \neq \{$родительный падеж$\} = A_2 - B_2$. A_1, B_1, A_2 и B_2 — здесь множества морфем в смысле Л. Ельмслева [38].

Все свойства оппозиций, пропорциональных влево, переносимы симметричным образом на оппозиции, пропорциональные вправо. Приводим формулировки трех теорем этого рода (их доказательство предоставляем читателю):

ТЕОРЕМА 1′. Если оппозиция (A_1/B_1) привативна в пользу B_1, а оппозиция (A_2/B_2) привативна в пользу B_2 и если оппозиция (A_1/B_1) пропорциональна вправо оппозиции (A_2B_2), тогда (A_1/B_1) пропорциональна оппозиции (A_2/B_2).

ТЕОРЕМА 2′. Если оппозиция (A_1B_1) привативна не в пользу B_1, а оппозиция (A_2/B_2) привативна не в пользу B_2, тогда оппозиция (A_1/B_1) пропорциональна оппозиции (A_2/B_2).

ТЕОРЕМА 3′. Если оппозиция (A_1/B_1) пропорциональна вправо оппозиции (A_2/B_2) и если оппозиция (A_2/B_2) привативна не в пользу B_2, тогда (A_1B_1) привативна не в пользу B_1.

ТЕОРЕМА 4. Если оппозиция (A_1/B_1) пропорциональна влево (вправо) оппозиции (A_2/B_2), тогда оппозиция (B_1/A_1) пропорциональна вправо (влево) оппозиции (B_2/A_2).

ДОКАЗАТЕЛЬСТВО. Доказательство вытекает из определений пропорциональности влево и пропорциональности вправо.

Попытаемся теперь доказать теорему следующего рода: «Если оппозиция (A_1/B_1) обладает свойством P, а оппозиция (A_2/B_2) пропорциональна оппозиции (A_1/B_1), то оппозиция (A_2/B_2) также обладает свойством P». Такое свойство P в этом случае мы будем называть инвариантом отношения пропорциональности, или инвариантом, сохраняющимся при пропорциональности.

ТЕОРЕМА 4′. Если свойство P есть инвариант отношения пропорциональности, тогда свойство не-P (т. е. свойство, полученное отрицанием P) также есть инвариант отношения пропорциональности.

ДОКАЗАТЕЛЬСТВО. Будем доказывать от противного. Если бы свойство не-P не было инвариантом, тогда существовала бы оппозиция (A_1/B_1), которая не обладает свойством P, но которая пропорциональна оппозиции (A_2/B_2), обладающей свойством P. Однако это противоречило бы предположению, что P есть инвариант по отношению пропорциональности.

ТЕОРЕМА 5. Если оппозиция (A_1/B_1) пропорциональна оппозиции (A_2/B_2), а (A_2/B_2) есть нулевая оппозиция, тогда (A_1B_1) также есть нулевая оппозиция.

ДОКАЗАТЕЛЬСТВО. Действительно, если $A_2 = B_2$ и если $A_1 - B_1 = A_2 - B_2$, то, следовательно, $A_1 \subseteq B_1$. Но из равенства $B_1 - A_1 = B_2 - A_2 = \varnothing$ вытекает, что $B_1 \subseteq A_1$, следовательно, $A_1 = B_1$.

ЗАМЕЧАНИЕ. Легко показать, что следующая теорема также верна. «Если (A_1/B_1) и (A_2/B_2) — это нулевые оппозиции, тогда оппозиция (A_1/B_1) пропорциональна оппозиции (A_2/B_2)». Действительно, $A_1 - B_1 = B_1 - A_1 = \varnothing$, $A_2 - B_2 = B_2 - A_2 = \varnothing$, откуда $A_1 - B_1 = A_2 - B_2$, $B_1 - A_1 = B_2 - A_2$.

ТЕОРЕМА 6. Если оппозиция (A_1/B_1) пропорциональна оппозиции (A_2/B_2), а (A_2/B_2) есть дизъюнктивная оппозиция, тогда оппозиция (A_1/B_1) эквиполентна или дизъюнктивна.

ДОКАЗАТЕЛЬСТВО. Действительно, $A_2 - B_2 = A_2$, $B_2 - A_2 = B_2$, откуда $A_1 - B_1 = A_2$, $B_1 - A_1 - B_2$, следовательно, $A_1 = B_1 \neq \varnothing \neq B_1 - A_1$, и поэтому тем более $A_1 \neq \varnothing \neq B_1$. Следовательно, если оппозиция (A_1/B_1) не эквиполентна, то она дизъюнктивна.

ТЕОРЕМА 7. Если оппозиция (A_1/B_1) пропорциональна оппозиции (A_2/B_2), а (A_2/B_2) есть эквиполентная оппозиция, тогда оппозиция (A_1/B_1) либо эквиполентна, либо дизъюнктивна.

ДОКАЗАТЕЛЬСТВО. По предположению $A_1 - B_1 = A_2 - B_2 \neq \varnothing$, $B_1 - A_1 = B_2 - A_1 = B_2 - A_2 \neq \varnothing$. Следовательно, если $A_1 \cap B_1 \neq \varnothing$, тогда оппозиция (A_1/B_1) дизъюнктивна, а если $A_1 \cap B_1 \neq \varnothing$, тогда оппозиция (A_1/B_1) эквиполентна.

Теоремы 6 и 7 можно обобщить в теореме 8.

ТЕОРЕМА 8. Свойство оппозиции быть дизъюнктивной или эквиполентной есть инвариант отношения пропорциональности.

ТЕОРЕМА 9. Если оппозиция (A_1/B_1) привативна в пользу (не в пользу) A_1, а оппозиция (A_2/B_2) пропорциональна оппозиции (A_1/B_1), тогда (A_2/B_2) привативна в пользу (не в пользу) A_2.

ДОКАЗАТЕЛЬСТВО. Пусть оппозиция (A_1B_1) привативна в пользу A_1. Из этого следует, что $B_1 - A_1 = \varnothing$, а $A_1 - B_1 \neq \varnothing$. По предположению, $B_2 - A_2 = B_1 - A_1$; следовательно, $B_2 - A_2 = \varnothing$. По предположению, $A_2 - B_2 = A_1 - B_1$; следовательно, $A_2 - B_2 \neq 0$ и оппозиция (A_2/B_2) привативна не в пользу A_2.

ЗАКЛЮЧЕНИЕ. Свойство оппозиции быть привативной есть инвариант отношения пропорциональности.

ТЕОРЕМА 10. Свойство оппозиции быть эквиполентной — не инвариант отношения пропорциональности.

ДОКАЗАТЕЛЬСТВО. Пусть $A_1 = \{$единственное число, именительный падеж, определенный статус$\}$, $B_1 = \{$множественное число, родительный падеж, неопределенный статус$\}$, $A_2 = \{$единственное число, именительный падеж, неопределенный статус, мужской род, сравнительная степень$\}$, $B_2 = \{$множественное число, родительный падеж, неопределенный статус, мужской род, сравнительная степень$\}$. Тогда $A_1 \cap B_1 = \varnothing$. Следовательно, оппозиция (A_1B_1) дизъюнктивна. $A_1 - B_1 = \{$един-

ственное число, именительный падеж, определенный статус} $= A_2 - B_2$, $B_1 - A_1 =$ {множественное число, родительный падеж, неопределенный статус} $= B_2 - A_2$. Следовательно, оппозиция (A_1/B_1) пропорциональна оппозиции (A_2/B_2). Однако оппозиция (A_2/B_2) не дизъюнктивна, а эквиполентна, так как $A_2 \cap B_2 =$ {мужской род, сравнительная степень} $\neq \varnothing$.

Из этого доказательства очевидна справедливость и следующей теоремы:

ТЕОРЕМА 11. Свойство оппозиции быть дизъюнктивной не есть инвариант отношения пропорциональности.

ТЕОРЕМА 12. Свойство оппозиции быть собственной не есть инвариант отношения пропорциональности.

ДОКАЗАТЕЛЬСТВО. Пусть $A_1 =$ {именительный падеж}, $B_1 =$ {единственное число, именительный падеж}, $A_2 = \varnothing$, $B_2 =$ {единственное число}. Отсюда $A_1 - B_1 = \varnothing = A_2 - B_2$, $B_1 - A_1 =$ {единственное число} $= B_2 - A_2$. Следовательно, оппозиция (A_1/B_1) пропорциональна оппозиции (A_2/B_2). Но (A_2/B_2) — несобственная оппозиция, тогда как (A_1/B_1) — собственная оппозиция.

Введем теперь новый тип отношений между оппозициями, а именно отношение гомогенности. Оппозиция (A_1/B_1) гомогенна с оппозицией (A_2/B_2), если $A_1 \cap B_1 = A_2 \cap B_2$, другими словами, если две оппозиции имеют одно и то же основание сравнения. Оппозиции (A_1/B_1) и (A_2/B_2) образуют гомогенную пару.

Если оппозиция (A/B) не гомогенна ни с какой другой оппозицией множества $\mathfrak{A}$, кроме (B/A), тогда мы говорим, что оппозиция (A/B) сингулярна в $\mathfrak{A}$.

ПРИМЕР 1. $A_1 =$ {красный, желтый}, $B_1 =$ {зеленый, желтый,} $A_2 =$ {белый, желтый}, $B_2 =$ {синий, желтый}. Отсюда $A_1 \cap B_1 =$ {желтый} $= A_2 \cap B_2$. Следовательно, (A_1/B_1) гомогенна с (A_2/B_2).

ПРИМЕР 2. $A_1 =$ {нейтральность в отношении звонкости, двугубность, фрикативность}, $B_1 =$ {глухость, губно-зубная артикуляция, фрикативность}, $A_2 =$ {звонкость, губно-зубная артикуляция, фрикативность}, $B_2 =$ {нейтральность в отношении звонкости, велярность, фрикативность}. $A_1 \cap B_1 =$ {фрикативность} $= A_2 \cap B_2$. Следовательно, оппозиция (A_1/B_1) гомогенна с (A_2/B_2). Иными словами, оппозиция румынских фонем *w* и *f* гомогенна оппозиции фонем *v* и *h*.

ПРИМЕР 3. $A_1 =$ {единственное число, именительный падеж, определенный статус}, $B_1 =$ {единственное число, родительный падеж, неопределенный статус}, $A_2 =$ {единственное число, именительный падеж, неопределенный статус}, $B_2 =$ {единственное число, родительный падеж, определенный статус}. Отсюда $A_1 \cap B_1 =$ {единственное число} $= A_2 \cap B_2$. Следовательно, оппозиция (A_1/B_1) гомогенна с (A_2/B_2).

ПРИМЕР 4. $\mathfrak{A} = \mathfrak{A}_\phi$. $A =$ {смычность, глухость, велярность}, $B =$ {смычность, звонкость, велярность}. Отсюда $A \cap B =$ {смычность, велярность}. Это основание сравнения не наблюдается в других оппозициях внутри множества $\mathfrak{A}_\phi$ (для румынского языка). Следовательно, (оппозиция (A/B) (т. е. между румынскими согласными фонемами *k* и *g*) сингулярна в $\mathfrak{A}_\phi$.

ПРИМЕР 5. $A_1 =$ {единственное число, именительный падеж, определенный статус}, $B_1 =$ {множественное число, именительный падеж, определенный статус}, $A_2 =$ {единственное число, именительный падеж, определенный статус, превосходная степень, мужской род}, $B_2 =$ {множественное число, именительный падеж, определенный статус, женский род, сравнительная степень}. Следовательно, $A_1 \cap B_1 =$ {именительный падеж, определенный статус} $= A_2 \cap B_2$, откуда следует, что оппозиция (A_1/B_1) гомогенна оппозиции (A_2/B_2). Из этого вытекает, что оппозиция (A_2/B_2) не сингулярна в $\mathfrak{A}_м$.

Рассмотрим теперь множество Γ, элементы которого суть множества, образованные грамматическими числами, падежами и статусами. Иными словами, Γ образовано множествами дифференциальных признаков румынских имен существительных. Обозначим через $\mathfrak{A}_{\text{сущ.}}$ множество оппозиций, которые образуют элементы Γ. Рассмотренная выше оппозиция между A_1 и A_2 сингулярна в $\mathfrak{A}_{\text{сущ.}}$, так как $A_1 - A_2 =$ $=$ {единственное число}, $A_2 - A_1 =$ {множественное число}, и не существует других чисел, кроме единственного и множественного.

Мы будем говорить, что две оппозиции (A_1/B_1) и (A_2/B_2) совпадают (или тождественны) и будем писать $(A_1/B_1) \equiv (A_2/B_2)$, если $A_1 = A_2$ и $B_1 = B_2$.

ТЕОРЕМА 12. Если оппозиция (A_1/B_1) пропорциональна и гомогенна с (A_2/B_2), то $(A_1/B_1) \equiv (A_2/B_2)$.

ДОКАЗАТЕЛЬСТВО. Из предположения о пропорциональности вытекает, что $A_1 - B_1 = A_2 - B_2$, $B_1 - A_1 = B_2 - A_2$. Из предположения о гомогенности вытекает, что $A_1 \cap B_1 = A_2 \cap B_2$. Но $A_1 = (A_1 \cap B_1) \cup$ $\cup (A_1 - B_1)$, $A_2 = (A_2 \cap B_2) \cup (A_2 - B_2)$, откуда $A_1 = (A_2 \cap B_2) \cup (A_2 -$ $- B_2) = A_2$. Аналогичным образом $B_1 = (B_1 \cap A_1) \cup (B_1 - A_1)$, $B_2 = (B_2 \cap$ $\cap A_2) \cup (B_2 - A_2)$, откуда $B_1 = B_2$.

ТЕОРЕМА 13. Если (A_1/B_1) — несобственная или дизъюнктивная оппозиция, а оппозиция (A_2/B_2) гомогенна с (A_1/B_1), то (A_2/B_2) также несобственная или дизъюнктивная оппозиция (другими словами, свойство оппозиции быть несобственной или дизъюнктивной есть инвариант отношения гомогенности).

ДОКАЗАТЕЛЬСТВО. Если (A_1/B_1) — дизъюнктивная оппозиция, то $A_1 \cap B_1 = \varnothing$. Из предположения о гомогенности вытекает, что $A_2 \cap B_2 =$ $= A_1 \cap B_1$, следовательно, $A_2 \cap B_2 = \varnothing$. Следовательно, если $A_2 \neq \varnothing \neq$ $\neq B_2$, то оппозиция (A_2/B_2) дизъюнктивна.

Если же (A_1/B_1) — несобственная оппозиция, то либо $A_1 = \varnothing$, либо $B_1 = \varnothing$. Предположим, например, что $A_1 = \varnothing$. В таком случае $A_1 \cap$ $\cap B_1 = 0$. Но оппозиция (A_2/B_2) гомогенна с (A_1/B_1), следовательно, $A_2 \cap B_2 = 0$, и поэтому оппозиция (A_2/B_2) либо собственна, либо дизъюнктивна.

ТЕОРЕМА 14. Свойство оппозиции быть привативной не есть инвариант отношения гомогенности.

Пусть $A_1 =$ {именительный падеж}, $B_1 =$ {единственное число, именительный падеж}, $A_2 =$ {множественное число, именительный падеж}, $B_2 = B_1$. Тогда $A_1 \cap B_1 =$ {именительный падеж} $= A_2 \cap B_2$, откуда оппозиция (A_1/B_1) гомогенна с (A_2/B_2). Но оппозиция (A_1/B_1) привативна, тогда как оппозиция (A_2/B_2) эквиполентна.

Из приведенного доказательства теоремы 14 вытекает также, что справедлива следующая теорема:

ТЕОРЕМА 15. Свойство оппозиции быть эквиполентной не есть инвариант отношения гомогенности.

Сведем в одну таблицу (табл. 2) перечень свойств, инвариантных по отношению пропорциональности или инвариантных по отношению гомогенности. Инвариантность будем обозначать знаком $+$, а отсутствие инвариантности знаком $-$. Оппозиция, привативная в пользу первого члена, будет называться привативной вправо, а привативная в пользу второго члена будет называться привативной влево.

Оппозиции, находящиеся в отношении гомогенности, соответствуют тому, что Н. С. Трубецкой называл многомерными оппозициями. Сингулярные оппозиции соответствуют двусторонним оппозициям Н. С. Трубецкого. Наименование «многомерная оппозиция» неудачно и может вызвать неправильное понимание, так как из этого названия

Таблица 2

Тип оппозиции	Инвариантность по отношению пропорциональности	Инвариантность по отношению гомогенности
1. Нулевая	+	—
2. Собственная	—	—
3. Привативная	+	—
4. Эквиполентная	—	—
5. Дизъюнктивная	—	—
6. Эквиполентная или дизъюнктивная	+	—
7. Несобственная или дизъюнктивная	—	+
8. Привативная влево	+	—
9. Привативная вправо . . .	+	—

неясно, что речь идет не о типе оппозиции, а о типе отношений между оппозициями.

Дизъюнктивные оппозиции Н. С. Трубецкого в явной форме не рассмотрены. Он включил их в класс эквиполентных оппозиций, определив последние как оппозиции с непустыми разностными множествами. Дизъюнктивные оппозиции были рассмотрены у Ж. Кантино, который назвал их отношением внешности. Как указано выше, Ж. Кантино предложил использовать термин «отношение» вместо «оппозиция». Наиболее разумным представляется использование обоих терминов, иначе пришлось бы говорить не об «отношениях между оппозициями», а об «отношениях между отношениями» или об «оппозициях оппозиций».

Термины «привативная оппозиция», «основание сравнения оппозиций» употребляются у Н. С. Трубецкого в смысле, сходном с нашим.

Объединение разностных множеств одной оппозиции соответствует тому, что Н. С. Трубецкой называет «характеристикой оппозиции». Если дана оппозиция (A/B), то характеристика ее — $(A—B) \cup (B—A)$. Это выражение, образованное множествами A и B, называется в теории множеств симметричной разностью множеств A и B и обозначается $A \triangle B$.

Вслед за Н. С. Трубецким примем для множества $A \triangle B$ наименование характеристики оппозиции (A/B).

Как вытекает из табл. 2, отношение пропорциональности допускает гораздо больше инвариантов, чем отношение гомогенности. Благодаря этому отношение пропорциональности играет несравненно бо́льшую роль, чем отношение гомогенности. Как видно из изложенного, тип эквиполентных оппозиций и тип дизъюнктивных оппозиций каждый порознь не являются инвариантами отношений пропорциональности, но тип, представляющий собой логическую сумму этих двух типов, оказывается таким инвариантом (табл. 2, строка 6). Это вполне соответствует концепции Н. С. Трубецкого. Н. С. Трубецкой не рассматривает типы оппозиций, помещенные в строках 4 и 5 табл. 2, но зато рассматривает тип оппозиций, помещенный у нас в строке 6; этот тип у него назван типом «эквиполентных оппозиций». А если учесть, что Н. С. Трубецкой не рассматривал также типов оппозиций, которые мы поместили в строках 2 и 7, получается, что у Н. С. Трубецкого

было прекрасное интуитивное понимание роли инвариантов. И действительно, он рассматривал только типы оппозиций, инвариантные по отношению пропорциональности, хотя делал это неосознанно и в неявной форме.

7. Отношения эквивалентности

Постараемся показать, что некоторые из типов отношений, изученных до сих пор, входят, несмотря на их различие, в общую схему.

Рассмотрим сначала отношение равенства между множествами. Каждое множество A равно самому себе: $A=A$. Действительно, если $x \in A$, то $x \in A$, откуда $A \subseteq A$, следовательно, $A = A$. Мы говорим поэтому, что отношение равенства рефлексивно. Если $A = B$, то $B = A$, так как из равенства $A = B$ следует, что $A \subseteq B$ и $B \subseteq A$, следовательно, $B = A$; мы говорим, что отношение равенства симметрично. Если $A = B$ и $B = C$, то $A = C$; действительно, из равенств $A = B$ и $B = C$ следует, что $A \subseteq B$ и $B \subseteq C$, откуда $A \subseteq C$; с другой стороны, из равенств $C = B$ и $B = A$ следует, что $C \subseteq B$ и $B \subseteq A$, откуда $C \subseteq A$. А если $A \subseteq C$ и $C \subseteq A$, то, следовательно, $A = C$. Мы говорим, что отношение равенства транзитивно.

Пусть $\mathfrak{A}$ есть множество оппозиций. Каждая оппозиция (A/B) пропорциональна себе самой: $(A/B) \sim (A/B)$. Это есть свойство рефлексивности отношения пропорциональности. Если $(A_1/B_1) \sim (A_2/B_2)$, то $(A_2/B_2) \sim (A_1/B_1)$. Действительно, благодаря симметричности отношения равенства из равенства $A_1—B_1 = A_2—B_2$ следует, что $A_2—B_2 = A_1—B_1$, а из равенства $B_1—A_1 = B_2—A_2$ вытекает, что $B_2—A_2 = B_1—A_1$. Таким образом, мы установили, что отношение пропорциональности симметрично. Наконец, можно легко показать, пользуясь транзитивностью отношения равенства, что если $(A_1/B_1) \sim (A_2/B_2)$ и $(A_2/B_2) \sim (A_3/B_3)$, то $(A_1/B_1) \sim (A_3/B_3)$. Действительно, из равенств $A_1—B_1 = A_2—B_2$ и $A_2—B_2 = A_3—B_3$ следует, что $A_1—B_1 = A_3—B_3$; из равенств же $B_1—A_1 = B_2—A_2$ и $B_2—A_2 = B_3—A_3$ следует, что $B_1—A_1 = B_3—A_3$. Мы установили, таким образом, транзитивность отношения пропорциональности.

Аналогичным образом можно доказать, что отношение гомогенности между оппозициями обладает теми же свойствами рефлексивности, симметричности и транзитивности. Каждая оппозиция гомогенна сама с собой; если оппозиция (A_1/B_1) гомогенна с (A_2/B_2), то оппозиция (A_2/B_2) гомогенна с (A_1/B_1); если оппозиция (A_1/B_1) гомогенна с (A_2/B_2), а оппозиция (A_2/B_2) гомогенна с (A_3/B_3), то оппозиция (A_1/B_1) гомогенна с (A_3/B_3).

Показанные выше три свойства отношения равенства, отношения пропорциональности и отношения гомогенности могут быть обобщены следующим образом:

Пусть R есть определенное отношение между элементами множества $\mathfrak{M}$. Будем писать xRy тогда и только тогда, когда элемент x находится в отношении R к y. Предположим, что имеют место следующие три свойства:

1) для каждого $x \in \mathfrak{M}$ верно, что xRx (свойство рефлексивности);

2) если $x \in \mathfrak{M}$, $y \in \mathfrak{M}$ и xRy, то yRx (свойство симметричности);

3) если $x \in \mathfrak{M}$, $y \in \mathfrak{M}$, $z \in \mathfrak{M}$, xRy и yRz, то xRz (свойство транзитивности).

Всякий раз, когда отношение R, определенное между элементами множества $\mathfrak{M}$, обладает свойствами (1), (2) и (3), мы будем говорить, что R есть отношение эквивалентности в $\mathfrak{M}$. Если xRy, мы говорим, что x эквивалентен y.

На основе показанного выше мы можем утверждать следующее: отношение равенства есть отношение, эквивалентное для любого множества множеств;

отношение пропорциональности есть отношение эквивалентности для любого множества оппозиций;

отношение гомогенности есть отношение эквивалентности для любого множества оппозиций.

В математике доказывается следующая теорема (см., например, [45]):

ТЕОРЕМА 16. Пусть R — отношение эквивалентности в $\mathfrak{M}$. Из множества $\mathfrak{M}$ выделяются подмножества со следующими свойствами: 1) каждый элемент $\mathfrak{M}$ входит в одно из этих подмножеств и только в одно, 2) если x и y входят в одно и то же подмножество, то xRy; 3) если x и y входят в разные подмножества, то x не находится в отношении R к y.

Подмножества множества $\mathfrak{M}$, которые упоминаются в этой теореме, называются классами эквивалентности. Класс эквивалентности по отношению R, по определению, образован всеми элементами $\mathfrak{M}$, эквивалентными данному элементу. Чтобы показать, что эти классы зависят от выбора отношения R, будем называть их классами R-эквивалентности. В частном случае, где $\mathfrak{M}$ есть множество оппозиций, а R — отношение пропорциональности, классы эквивалентности будем называть классами пропорциональных оппозиций, или, короче, классами пропорциональности. Если R есть отношение гомогенности, мы получаем классы гомогенных оппозиций, или, короче, классы гомогенности. Из последней теоремы вытекает, что оппозиция принадлежит только одному классу пропорциональности и только одному классу гомогенности. Из того факта, что две разные оппозиции не могут быть в одно и то же время пропорциональными друг другу и гомогенными друг с другом, следует, что оппозиция полностью определяется указанием класса пропорциональности и класса гомогенности, в которые она входит (см. в связи с этим теорему 12 настоящей работы).

Интересно установить не только сходства оппозиций между собой, но и их различия. С этой целью докажем теорему 17.

ТЕОРЕМА 17. Если оппозиция (A_1/B_1) пропорциональна оппозиции (A_2/B_2), то имеет место одна из следующих двух ситуаций:

1) $A_1 = A_2$, $B_1 = B_2$, 2) $A_1 \neq A_2$, $B_1 \neq B_2$.

ДОКАЗАТЕЛЬСТВО. Ситуация (1) имеет место тогда, когда две оппозиции совпадают. Покажем, что всякий раз, когда оппозиции не совпадают, имеет место ситуация (2). Действительно, рассмотрим две пропорциональные оппозиции, имеющие в качестве первого члена множество A.

$(A/B_1) \sim (A/B_2)$. Тогда

$$A - B_1 = A - B_2, \tag{1}$$

$$B_1 - A = B_2 - A. \tag{2}$$

Из (1) вытекает, что $A - (A - B_1) = A - (A - B_2)$. Но $A - (A - B_1) = B_1 \cap A$, $A - (A - B_2) = B_2 \cap A$, следовательно,

$$B_1 \cap A = B_2 \cap A. \tag{3}$$

Поскольку $B_1 = (B_1 - A) \cup (B_1 \cap A)$, а $B_2 = (B_2 - A) \cup (B_2 \cap A)$, то из (2) и (3) следует, что $B_1 = B_2$.

Следовательно, невозможно, чтобы две пропорциональные оппозиции, отличаясь одним членом, не отличались другим членом.

ТЕОРЕМА 18. Множество не может участвовать в двух разных оппозициях, пропорциональных друг другу.

ДОКАЗАТЕЛЬСТВО. Эта теорема отличается от предыдущей только тем, что здесь равенство $B_1 = B_2$ должно вытекать и из пропорциональности $(A/B_1) \sim (B_2/A)$. Чтобы доказать эту импликацию, достаточно в эксплицитной форме написать отношение пропорциональности:

$$A - B_1 = B_2 - A,$$
$$B_1 - A = A - B_2.$$

Эти равенства возможны лишь в случае, если $A - B_1 = B_2 - A = B_1 - A = A - B_2 = \varnothing$. Действительно, если бы одна из этих разностей, например $A - B_1$, была бы непуста, тогда из $x \in A - B_1$ вытекало бы, что $x \in A$, откуда $x \notin B_2 - A$, что противоречит первому равенству $(A - B_1 = B_2 - A)$. Следовательно, $A = B_1 = B_2$. Теорема доказана.

Понятия «класс пропорциональных оппозиций» и «класс гомогенных оппозиций» уточняют и обобщают некоторые понятия, рассмотренные Н. С. Трубецким, например, понятие «корреляция» (= «класс пропорциональных и привативных оппозиций») или понятие «ряд пропорциональных оппозиций», а также некоторые понятия, введенные Ж. Кантино, например, понятие «цепочка гомогенных оппозиций».

Теоремы, касающиеся инвариантности некоторых типов оппозиций по отношению пропорциональности или по отношению гомогенности, теперь можно выразить так, как это показано на табл. 2.

Если класс пропорциональности содержит привативную оппозицию, то всякая оппозиция этого класса привативна; если класс пропорциональности содержит нулевую оппозицию, то всякая оппозиция этого класса нулевая; если класс пропорциональности содержит эквиполентную или дизъюнктивную оппозицию, тогда каждая оппозиция этого класса либо эквиполентна, либо дизъюнктивна и т. д.

Из этого следует, что среди различных классов пропорциональности наиболее интересны те, которые образованы оппозициями, по структуре своей инвариантными по отношению пропорциональности. В соответствии с табл. 2 это следующие классы: а) класс нулевых оппозиций, б) класс привативных оппозиций, в) класс эквиполентных или дизъюнктивных оппозиций, г) класс оппозиций, привативных влево, д) класс оппозиций, привативных вправо.

Таким образом, классы пропорциональности типов (а), (в), (г) и (д), оказывается, имеют то же право на независимое существование, как и класс типа (б), на который обратил особое внимание Н. С. Трубецкой, присвоив ему название «корреляция». Этот факт не мог остаться незамеченным: позднее А. Мартине [40] распространил термин «корреляция» и на классы пропорциональности типа (в). (Как и Н. С. Трубецкой, А. Мартине использует термин «эквиполентный» в том смысле, в каком мы говорим «эквиполентный или дизъюнктивный»). Что касается класса пропорциональности типа (а), он является тривиальным не только потому, что образован нулевыми оппозициями, но и потому, что две нулевые оппозиции всегда пропорциональны друг другу, и поэтому все нулевые оппозиции образуют один класс пропорциональности.

Классы же типов (г) и (д) не были в явной форме рассмотрены лингвистами.

Имеет смысл сохранить идею определения А. Мартине, распространив это определение на другие классы оппозиций, инвариантных по отношению пропорциональности. Будем поэтому называть корреляцией всякий класс пропорциональности типов (а), (б), (в), (г) или (д). Итак, корреляцией будем называть класс пропорциональности, образованный оппозицией, имеющей инвариантную структуру по отношению пропорциональности. В этом смысле каждый тип оппозиции, инвариант-

ный по отношению пропорциональности, порождает тип корреляции. Следовательно, корреляция — это частный случай класса пропорциональности, однако только этот частный случай интересен с научной точки зрения. Действительно, как видно из табл. 2, из оппозиций структуры, неинвариантной по отношению пропорциональности, образованы такие классы пропорциональности, как (е) класс собственных оппозиций, (ж) класс эквиполентных оппозиций, (з) класс дизъюнктивных оппозиций, (и) класс собственных или дизъюнктивных оппозиций. Классы типа (е), (ж), (з) или (и) имеют разнородный состав, внутренняя структура оппозиций, составляющих любой такой класс, не сохраняется благодаря отношению пропорциональности. В результате эти классы представляют меньший научный интерес.

Другая важная иллюстрация теоремы 16 следующая. Пусть A есть множество слов и пусть X есть множество контекстов (понятие «контекст» мы уточним в другой нашей работе [47]). Сопоставим каждому слову $a \in A$ определенную часть $\mathfrak{e}(a)$ множества X, т. е. некоторое подмножество контекстов (в содержательной интерпретации множество $\mathfrak{e}(a)$ было бы множеством контекстов, в которых появляется слово a). Мы будем говорить, что два слова a и b одинаково распределены, если $\mathfrak{e}(a) = \mathfrak{e}(b)$. Поскольку отношение одинаковости распределения восходит к отношению равенства двух множеств, это отношение есть отношение эквивалентности. Классы эквивалентности, соответствующие ей по теореме 16, это так называемые классы дистрибуции (классы распределения) или формальные классы, рассматриваемые в дескриптивной лингвистике. Класс дистрибуции слова a есть множество слов, которые появляются в тех и только в тех контекстах, в каких появляется слово a.

Мы видим, что три лингвистических понятия из числа наиболее важных — ряды пропорциональных оппозиций, цепочки гомогенных оппозиций и классы дистрибуции — все восходят к одной и той же логической процедуре: разложению множества на классы R-эквивалентности. Эти три понятия имеют различное лингвистическое содержание, что проявляется в различном выборе исходного множества и в выборе отношения R. Но как только мы абстрагируемся от природы множеств и отношения R, эти три понятия становятся неразличимы. Эта тождественность структуры отражает часть того, что некоторые лингвисты называют инвариантной структурой языка. Но точный смысл этого явления обнаружится только тогда, когда мы дадим строгую формулировку так называемому лингвистическому изоморфизму.

Отношения эквивалентности обнаруживаются в определении некоторых основных понятий лингвистики, например фонемы, морфемы, падежа, части речи и т. д. См., например, [19], [10], [14], [15] и [6].

Введем теперь понятие, которое интересно само по себе и которое, кроме того, позволит нам выразить важное следствие из теоремы 16.

Дано множество A и подмножества $A_1, A_2, \ldots, A_n \ldots$ множества A, предполагаемые взаимно не пересекающимися. Другими словами, ни один элемент множества A не принадлежит одновременно двум из указанных подмножеств. Предположим еще, что

$$A = A_1 \cup A_2 \cup \ldots \cup A_n \cup \ldots \tag{4}$$

Тогда мы говорим, что отношение (4) определяет разбиение множества A.

Если в правой части равенства (4) есть конечное число членов, тогда мы говорим, что разбиение конечно.

С помощью понятия разбиения выведем из теоремы 16 важное следствие.

СЛЕДСТВИЕ: Если R — отношение эквивалентности в множестве $\mathfrak{M}$, то классы R-эквивалентности определяют разбиение множества $\mathfrak{M}$.

Возможности использования изученных в математике отношений между множествами значительно более разнообразны, нежели те, которые можно видеть на приведенных примерах. Так, например, различные типы дистрибуций, изучаемые в современной лингвистике (см., например, [3]), могут быть описаны с помощью отношения включения и операции пересечения. Рассмотрим множество A слов и множество X контекстов. Сопоставим каждому слову $a \in A$ определенную часть $\mathfrak{e}(a) \subset X$ (в содержательной интерпретации $\mathfrak{e}(a)$ есть множество контекстов, в которых может появиться слово a). Мы можем говорить, что два слова a и b находятся в дополнительном распределении, если соответствующие множества $\mathfrak{e}(a)$ и $\mathfrak{e}(b)$ находятся в дизъюнктивной оппозиции, другими словами, если $\mathfrak{e}(a) \cap \mathfrak{e}(b) = \varnothing$. Если два слова a и b не находятся в отношении дополнительного распределения, т. е. если $\mathfrak{e}(a) \cap \mathfrak{e}(b) \neq \varnothing$, то мы говорим, что a и b находятся в контрастном распределении. Кроме того, мы можем, вслед за Ф. де Соссюром, говорить, что a и b находятся в ассоциативном отношении. Контрастные распределения могут быть, в свою очередь, разного рода: тождественные, дефективные и эквиполентные. Если $\mathfrak{e}(a) = \mathfrak{e}(b)$, другими словами, если $\mathfrak{e}(a)$ и $\mathfrak{e}(b)$ находятся в нулевой оппозиции, то мы говорим, что a и b находятся в тождественном распределении или что они одинаково распределены. Эта ситуация ведет к выделению классов дистрибуции, которые, как было показано выше, являются классами эквивалентности по отношению тождественного распределения. Если $\mathfrak{e}(a) \subset \mathfrak{e}(b)$ или $\mathfrak{e}(b) \subset \mathfrak{e}(a)$, другими словами, если оппозиция между $\mathfrak{e}(a)$ и $\mathfrak{e}(b)$ привативна, тогда мы говорим, что a и b находятся в дефективном распределении. Дефективное распределение бывает в пользу a или в пользу b, в зависимости от того, является ли оппозиция привативной в пользу $\mathfrak{e}(a)$ или в пользу $\mathfrak{e}(b)$. Наконец, если оппозиция между $\mathfrak{e}(a)$ и $\mathfrak{e}(b)$ эквиполентна, то мы говорим, что a и b находятся в эквиполентном распределении.

Следует заметить, что множества, с помощью которых мы описывали различные типы распределения, образованы элементами, значительно более сложными, чем те, которые участвовали в описании оппозиций между фонемами. Фонема есть множество неразложимых элементов (звонкость, зубная артикуляция, смычность и др.), т. е. таких элементов, которые нельзя разложить на более простые компоненты, тогда как контекст есть сложная единица, которая, чтобы быть изученной, должна быть разложена на более простые элементы. По этой причине изучение типов распределения, основанное на классификации оппозиций между множествами контекстов, не достигло таких успехов, как изучение фонем посредством оппозиций между множествами дифференциальных фонологических признаков. Оппозиции между множествами тем более интересны, чем более рудиментарный и, если можно, сводимый к более простым компонентам характер носят элементы, из которых образованы множества.

Несводимые элементы мы будем с лингвистической точки зрения называть лингвистическими квантами, а процесс разложения более сложных образований на такие простейшие элементы будем называть процессом квантования. Там, где это невозможно, предпочтительнее использование других типов отношений, изучение которых выходит за рамки данной статьи.

Для количественного изучения двоичных и троичных отношений между множествами весьма полезно обращение к работе [5].

II. УПОРЯДОЧЕННЫЕ ОППОЗИЦИИ, ПАРАДИГМЫ, МОРФЕМЫ И КВАЗИМОРФЕМЫ

1. Введение

Теория лингвистических оппозиций, развитая Н. С. Трубецким и Ж. Кантино, касается оппозиции между множествами лингвистических элементов. Такая теория пригодна для изучения, например, оппозиций между фонемами, рассматриваемыми как множества дифференциальных признаков, или оппозиций между множествами морфем в смысле Ельмслева. Но если мы хотим изучать оппозицию между словами, рассматриваемыми как множество абстрактных звуков (в смысле С. К. Шаумяна — [18] и [21]), мы найдем, что русские слова *от* и *то* находятся в нулевой оппозиции, так как каждое из них сводится к множеству $\{o, m\}$. Интуитивно очевидно, что *от* и *то* — разные слова, но теория лингвистических оппозиций пока не сделала неизбежного вывода из этого: необходимо точно определить и исследовать оппозиции между упорядоченными множествами языковых элементов. В первом разделе работы мы коснулись логического аспекта оппозиций между неупорядоченными множествами. Теперь мы обратимся, не входя в детали, к оппозициям между упорядоченными множествами. В результате мы получим логическую модель понятия парадигмы и, вследствие этого, понятия морфемы.

2. Цепочки и упорядоченные множества

Пусть E есть множество элементов любой природы. Любой конечный ряд элементов E (каждый элемент может встречаться в ряду несколько раз) будем называть цепочкой. Будем считать цепочкой также пустую цепочку, не содержащую ни одного элемента. Если x, y, z,... есть элементы E, тогда xyz, xyx, yxx, zzz — это различные цепочки этих элементов. Существуют цепочки, различающиеся по составу элементов (например, xyz и yxx), тогда как другие цепочки различаются лишь порядком элементов (например, xyx и yxx). Следовательно, цепочка — это конечное и упорядоченное множество элементов множества E.

Подмножество β цепочки α есть подцепочка цепочки α, если элементы подмножества β располагаются в том же порядке, что и в α, и если два элемента подмножества β являются смежными в β тогда и только тогда, когда они смежны в α. Например, xy и yz — это подцепочки цепочки xyz, но xz не есть подцепочка цепочки xyz, так как x и z не смежны в xyz.

В каждой цепочке будем выделять начальный элемент и конечный элемент цепочки. Подцепочка цепочки α, имеющая тот же начальный элемент, что и α, называется начальной подцепочкой цепочки α; подцепочка цепочки α, имеющая тот же конечный элемент, что и α, называется конечной подцепочкой цепочки α. Подцепочка, не являющаяся ни начальной, ни конечной, называется срединной подцепочкой.

Упорядоченная оппозиция есть упорядоченная пара цепочек. Возьмем две цепочки λ и μ. Оппозиция между λ и μ будет обозначена посредством (λ, μ). Наибольшая начальная подцепочка, общая цепочкам λ и μ, называется начальным основанием оппозиции (λ, μ). Чтобы указать, что оппозиция (λ, μ) рассматривается в отношении своего начального основания, мы будем говорить, что оппозиция (λ, μ) есть начальная оппозиция, и будем обозначать ее через $(\lambda, \mu)_i$. Мы можем ввести также понятие конечной оппозиции. Наибольшая конечная подцепочка, общая цепочкам λ и μ, называется конеч-

ным основанием оппозиции (λ, μ). Чтобы указать, что оппозиция (λ, μ) будет рассматриваться в отношении своего конечного основания, мы будем говорить, что оппозиция (λ, μ) есть конечная оппозиция, и будем обозначать ее через $(\lambda, \mu)_f$.

Мы могли бы также ввести понятие срединной оппозиции, но в дальнейшем мы будем оперировать только начальными и конечными оппозициями. Мы будем детально исследовать в эксплицитной форме только начальные оппозиции; те же рассуждения можно симметричным образом перенести и на конечные оппозиции. Конечные и начальные упорядоченные оппозиции будут называться крайними оппозициями.

3. Классификация начальных оппозиций

Определим операцию соединения цепочек. Если даны две цепочки λ и μ, операция соединения приведет к созданию цепочки $\lambda\mu$, в которой цепочка μ приложена справа к цепочке λ. Операция соединения ассоциативна, но не коммутативна (легко видеть, что операцию соединения можно определить для любого конечного числа цепочек). Например, если $\lambda = xyz$ и $\mu = acb$, то $\lambda\mu = xyzacb$.

Рассмотрим начальную оппозицию $(\lambda, \mu)_i$. Пусть ν есть ее начальное основание. Существует единственная цепочка λ' и единственная цепочка μ', такие, что $\lambda = \nu\lambda'$ и $\mu = \nu\mu'$. Цепочки λ' μ' называются дифференциальными подцепочками оппозиции $(\lambda, \mu)_i$.

Оппозицию $(\lambda, \mu)_i$ будем называть нулевой оппозицией, если $\lambda = \mu$, т. е. если дифференциальные подцепочки пусты.

Оппозицию $(\lambda, \mu)_i$ будем называть привативной слева, или привативной в пользу μ, если λ' есть пустая цепочка, а цепочка μ' не пуста. Оппозиция $(\lambda, \mu)_i$ будет называться привативной справа, или привативной в пользу λ, если μ' есть пустая цепочка, а цепочка λ' не пуста. Оппозиция $(\lambda, \mu)_i$, привативная слева или справа, называется привативной оппозицией. Пример: если $\lambda = стол$, $\mu = стола$, оппозиция $(\lambda, \mu)_i$ привативна слева. Если $\lambda = стола$, $\mu = стол$, оппозиция $(\lambda, \mu)_i$ привативна справа.

Оппозицию $(\lambda, \mu)_i$ будем называть дизъюнктивной, если ее начальное основание есть пустая цепочка. Например, если $\lambda = город$ и $\mu = книга$, оппозиция $(\lambda, \mu)_i$ дизъюнктивна; если $\lambda = до$, $\mu = то$, оппозиция $(\lambda, \mu)_i$ также дизъюнктивна.

Оппозиция $(\lambda, \mu)_i$ называется эквиполентной, если она не нулевая, не привативная и не дизъюнктивная. Например, если $\lambda = стола$, а $\mu = столу$, то оппозиция $(\lambda, \mu)_i$ эквиполентна, так как тогда $\nu = стол$, $\lambda' = -а$ и $\mu' = -у$; если $\lambda = бумага$, а $\mu = быть$, то оппозиция $(\lambda, \mu)_i$ также эквиполентна, так как в этом случае $\nu = б-$, $\lambda' = умага$, а $\mu' = -ыть$.

4. Отношения между начальными оппозициями

Две начальные оппозиции $(\lambda_1, \mu_1)_i$ и $(\lambda_2, \mu_2)_i$ будем называть гомогенными, если у них есть одно и то же начальное основание, т. е. если $\nu_1 = \nu_2$, где ν_1 есть начальное основание оппозиции $(\lambda_1, \mu_1)_i$, а ν_2 есть начальное основание оппозиции $(\lambda_2, \mu_2)_i$. Например, если $\lambda_1 = стол- = \lambda_2$, $\mu_1 = стола$, $\mu_2 = столу$, то в этом случае $\nu_1 = стол- = \nu_2$, т. е. оппозиции $(\lambda_1, \mu_1)_i$ и $(\lambda_2, \mu_2)_i$ гомогенны.

Легко видеть, что отношение гомогенности между начальными оппозициями есть отношение эквивалентности. Также легко показать, что если из двух однородных оппозиций $(\lambda_1, \mu_1)_i$ и $(\lambda_2, \mu_2)_i$ оппозиция $(\lambda_1, \mu_1)_i$ дизъюнктивна, то и оппозиция $(\lambda_2, \mu)_i$ также дизъюнктивна.

Если не существует начальных оппозиций, гомогенных оппозиций $(\lambda, \mu)_i$, мы говорим, что $(\lambda, \mu)_i$ есть сингулярная оппозиция.

Две начальные оппозиции $(\lambda_1, \mu_1)_i$ и $(\lambda_2, \mu_2)_i$ называются **пропорциональными**, если у них те же дифференциальные подцепочки, т. е. если $\lambda'_1 = \lambda'_2$ и $\mu'_1 = \mu'_2$, где λ'_1 и μ'_1 есть дифференциальные подцепочки оппозиции $(\lambda_1, \mu_1)_i$, а λ'_2 и μ'_2 — это дифференциальные подцепочки оппозиции $(\lambda_2, \mu_2)_i$. Например, если $\lambda_1 = дом$, $\lambda_2 = луг$, $\mu_1 = дома́$, а $\mu_2 = луга́$, то $\lambda'_1 = \lambda'_2 = 0$, $\mu'_1 = \mu'_2 = -а$, откуда оппозиция $(\lambda_1, \mu_1)_i$ пропорциональна оппозиции $(\lambda_2, \mu_2)_i$.

Легко видеть, что отношение пропорциональности начальных оппозиций есть отношение эквивалентности. Также легко показать, что, когда даны две пропорциональных оппозиции $(\lambda_1, \mu_1)_i$ и $(\lambda_2, \mu_2)_i$, то в случае, если $(\lambda_1, \mu_1)_i$ есть нулевая оппозиция, $(\lambda_2, \mu_2)_i$ — также нулевая оппозиция, если оппозиция $(\lambda_1, \mu_1)_i$ привативна слева, оппозиция $(\lambda_2, \mu_2)_i$ также привативна слева, если $(\lambda_1, \mu_1)_i$ привативна справа, $(\lambda_2, \mu_2)_i$ также привативна справа, если же оппозиция $(\lambda_1, \mu_1)_i$ эквиполентна или дизъюнктивна, то и оппозиция $(\lambda_2, \mu_2)_i$ также эквиполентна или дизъюнктивна. Но эквиполентность оппозиции $(\lambda_1\mu_1)_i$ не обязательно влечет за собой эквиполентность оппозиции $(\lambda_2, \mu_2)_i$ и дизъюнктивность оппозиции $(\lambda_1\mu_1)_i$ не влечет за собой дизъюнктивности оппозиции $(\lambda_2, \mu_2)_i$. Доказательство этих положений аналогично доказательству соответствующих теорем, данному для оппозиции между неупорядоченными множествами.

5. Конечные оппозиции

Пусть λ и μ — две цепочки и пусть ω — наибольшая конечная подцепочка, общая цепочкам λ и μ. Оппозиция (λ, μ), рассмотренная в отношении подцепочки ω, является, по определению, **конечной оппозицией** и будет обозначаться через $(\lambda, \mu)_f$.

Существует единственная цепочка λ'' и единственная цепочка μ'', такие, что $\lambda = \lambda''\omega$ и $\mu = \mu''\omega$. Такие цепочки λ'' и μ'' называются **дифференциальными подцепочками** оппозиции $(\lambda, \mu)_f$. Оппозиция $(\lambda, \mu)_f$ называется **нулевой**, если $\lambda = \mu$, **привативной слева**, если $\lambda'' = 0$ и $\mu'' \neq 0$, **привативной справа**, если $\mu'' = 0$ и $\lambda'' \neq 0$, **привативной**, если она привативна слева или справа, **дизъюнктивной**, если $\omega = 0$, **эквиполентной**, если она не привативная, не нулевая и не дизъюнктивная. Примеры: если $\lambda = от$ и $\mu = вот$, то оппозиция $(\lambda, \mu)_f$ привативна слева, а оппозиция $(\mu, \lambda)_f$ привативна справа. Действительно, в обоих случаях $\omega = от$, $\lambda'' = 0$, $\mu'' = в$-. Если $\lambda = мужчины$ и $\mu = столы$, то оппозиция $(\lambda, \mu)_f$ эквиполентна, так как $\omega = $-ы, и поэтому оппозиция $(\lambda, \mu)_f$ не дизъюнктивна, а $\lambda'' = мужчин$-, $\mu'' = стол$-, и поэтому оппозиция $(\lambda, \mu)_f$ не нулевая и не привативная.

6. Отношения между конечными оппозициями

Две конечные оппозиции $(\lambda_1, \mu_1)_f$ и $(\lambda_2, \mu_2)_f$ называются **гомогенными**, если у них то же конечное основание. Например, если $\lambda_1 = но$, $\mu_1 = \lambda_2 = то$, $\mu_2 = до$, то $(\lambda_1, \mu_1)_f$ и $(\lambda_2, \mu_2)_f$ имеют то же конечное основание: -о.

Отношение гомогенности между конечными оппозициями есть отношение эквивалентности. Если две конечные оппозиции гомогенны и если одна из них дизъюнктивна, то другая тоже дизъюнктивна. Если не существует конечных оппозиций, гомогенных оппозиции $(\lambda, \mu)_f$, то мы будем говорить, что $(\lambda, \mu)_f$ есть **сингулярная оппозиция**.

Две конечные оппозиции **пропорциональны**, если у них те же дифференциальные подцепочки. Например, если $\lambda_1 = но$, $\lambda_2 = на$, $\mu_1 = до$, $\mu_2 = да$, то оппозиции $(\lambda_1\mu_1)_f$ и $(\lambda_2, \mu_2)_f$ пропорциональны, так как $\lambda''_1 = \lambda''_2 = н$-, $\mu''_1 = \mu''_2 = д$-.

Отношение пропорциональности между конечными оппозициями есть отношение эквивалентности. Можно показать, что отношение пропорциональности между конечными оппозициями сохраняет все свойства отношения пропорциональности между начальными оппозициями (см. последнюю часть раздела 4).

7. Регулярная и парадигматическая морфология

Пусть E есть некоторое множество. Образуем цепочки из элементов множества E. Пусть $\mathfrak{E}$ есть некоторое множество таких цепочек. Назовем допустимыми словами все элементы множества $\mathfrak{E}$.

Пусть Π есть разбиение множества $\mathfrak{E}$.

$$\mathfrak{E} = \bigcup_j \mathfrak{E}_j$$

Пусть $\mathfrak{F}$ есть некоторое множество цепочек, элементы которых принадлежат множеству $\mathfrak{E}$. Будем называть языком триаду $\{\mathfrak{E}, \Pi, \mathfrak{F}\}$; допустимыми фразами будем называть элементы множества $\mathfrak{F}$.

Член $\mathfrak{E}_i$ разбиения Π называется регулярным, если он обладает следующими свойствами:

Если $\lambda_1 \in \mathfrak{E}_j$, $\lambda_2 \in \mathfrak{E}_j$, $\mu_1 \in \mathfrak{E}_j$, $\mu_2 \in \mathfrak{E}_j$, то оппозиции $(\lambda_1 \mu_1)_i$ и $(\lambda_2, \mu_2)_i$ гомогенны.

Например, если $\mathfrak{E}_j$ есть совокупность форм единственного числа русского имени существительного *завод*, то $\mathfrak{E}_j$ есть регулярный член разбиения. Действительно, $\mathfrak{E}_j = \{$*завóд*, *завода*, *заводу*, *заводом*, *заводе*$\}$, и легко видеть, что все начальные оппозиции двух элементов множества $\mathfrak{E}_j$ имеют то же основание: *завод-*.

Если в языке $\mathfrak{L}$ все множества $\mathfrak{E}_j$ регулярны, то будем говорить, что такой язык $\mathfrak{L}$ имеет регулярную морфологию.

Пусть λ и μ — два допустимых слова, обладающих следующим свойством: не существует допустимой фразы, содержащей слово λ, такой, что при замене слова λ словом μ мы получили бы допустимую фразу. В таком случае мы говорим, что λ и μ находятся в отношении дополнительного распределения.

Пусть язык $\mathfrak{L} = (\mathfrak{E}, \Pi, \mathfrak{F})$ есть язык с регулярной морфологией. Мы будем называть множество $\mathfrak{E}_j$, получаемое при разбиении Π, парадигмой языка $\mathfrak{L}$, если для всяких $\lambda \in \mathfrak{E}_j$ и $\mu \in \mathfrak{E}_j$ λ и μ находятся в дополнительном распределении. Если всякое множество $\mathfrak{E}_j$ есть парадигма языка $\mathfrak{L}$, мы говорим, что язык $\mathfrak{L}$ имеет парадигматическую морфологию. Следовательно, каждый язык, имеющий парадигматическую морфологию, имеет регулярную морфологию, но обратное неверно.

Пусть $\mathfrak{L}$ есть некий язык. Мы будем говорить, что два множества $\mathfrak{E}_m$ и $\mathfrak{E}_n$, полученные при разбиении Π, гомологичны, если мы можем установить между множествами $\mathfrak{E}_m$ и $\mathfrak{E}_n$ взаимно-однозначное соответствие таким образом, что если $\lambda_m \in \mathfrak{E}_m$, $\mu_m \in \mathfrak{E}_m$, $\lambda_n \in \mathfrak{E}_n$, $\mu_n \in \mathfrak{E}_n$ и λ_n соответствует λ_m и μ_n соответствует μ_m, то оппозиции $(\lambda_m, \mu_m)_i$ и $(\lambda_n, \mu_n)_i$ пропорциональны. Легко видеть, что отношение гомологичности есть отношение эквивалентности.

Полученные таким образом классы эквивалентности будем называть классами гомологичности. В частности, в языке с парадигматической морфологией парадигмы распределяются в классы гомологичности.

Пусть $\mathfrak{E}_m$ и $\mathfrak{E}_n$ — это две гомологичные парадигмы языка, имеющего парадигматическую морфологию. Если $\lambda \in \mathfrak{E}_m$, $\mu \in \mathfrak{E}_m$, $\lambda' \in \mathfrak{E}_n$, $\mu' \in \mathfrak{E}_n$ и если λ' соответствует λ, а μ' соответствует μ, конечные оппозиции $(\lambda, \lambda')_f$ и $(\mu, \mu')_f$ пропорциональны. Это свойство является непосредственным следствием данных выше определений.

Приведем конкретный пример. Допустим, что $\mathfrak{E}$ есть лексика русского языка и для каждого $\lambda \in \mathfrak{E}$ множество $\mathfrak{E}_j$, содержащее λ, определяется как множество флективных форм λ. В этом случае, если $\mathfrak{E}_m$ есть множество флективных форм слова *завод*, а $\mathfrak{E}_n$ есть множество флективных форм слова *народ*, существует гомологичность между $\mathfrak{E}_m$ и $\mathfrak{E}_n$. Действительно, $\mathfrak{E}_m = \{$*завод*, *завода*, *заводу*, *заводом*, *заводе*, *заводы*, *заводов*, *заводам*, *заводами*, *заводах*$\}$, $\mathfrak{E}_n = \{$*народ*, *народа*, *народу*, *народом*, *народе*, *народы*, *народов*, *народам*, *народами*, *народах*$\}$. Если поставим в соответствие элементы того же ранга (по порядку следования) обоих множеств, условие пропорциональности будет выполнено, в чем легко убедиться. Например, оппозиция (*народа*, *народу*) пропорциональна оппозиции (*завода*, *заводу*)$_i$, оппозиция (*народами*, *народах*) пропорциональна оппозиции (*заводами*, *заводах*)$_i$ и т. д.

8. Морфемы и квазиморфемы

Пусть $\mathfrak{L}$ есть язык с парадигматической морфологией и пусть $\mathfrak{E}_1$ $\mathfrak{E}_2, \ldots, \mathfrak{E}_n$ суть парадигмы языка $\mathfrak{L}$. Цепочка λ, по определению, является морфемой языка $\mathfrak{L}$ по отношению к $\mathfrak{E}_j$, если λ есть основание или дифференциальная цепочка в начальной оппозиции двух членов множества $\mathfrak{E}_j$. Если λ есть основание, тогда эту морфему λ будем называть морфемой-основанием (основой) по отношению к $\mathfrak{E}_j$; если λ есть дифференциальная цепочка, тогда эта цепочка λ есть дифференциальная морфема по отношению к $\mathfrak{E}_j$. Отсюда ясно, что понятие морфемы определяется по отношению к некоторой парадигме.

Из определений, данных выше, вытекает:

1) в языке с парадигматической морфологией существует по отношению к каждой определенной парадигме только одна морфема-основание;

2) если две парадигмы гомологичны, в них те же дифференциальные цепочки.

Понятие морфемы было определено только для языков с парадигматической морфологией. Но естественные языки обычно не удовлетворяют этому условию. Можно для каждого языка определить понятие квазиморфемы; в частном случае языков с парадигматической морфологией понятие квазиморфемы сводится к понятию морфемы. Пусть язык $\mathfrak{L} = \{\mathfrak{E}, \Pi, \mathfrak{F}\}$ и пусть $\mathfrak{E}_j$ есть множество, полученное при разбиении Π. Всякая цепочка, являющаяся основанием или дифференциальной цепочкой начальной оппозиции между элементами множества $\mathfrak{E}_j$, есть, по определению, квазиморфема по отношению к $\mathfrak{E}_j$.

Существуют квазиморфемы-основания по отношению к $\mathfrak{E}_j$ и дифференциальные квазиморфемы по отношению к $\mathfrak{E}_j$. Возьмем пример из русского языка. Если $\mathfrak{E}_j$ есть совокупность флективных форм слова *завод*, то $\mathfrak{E}_j$ не есть регулярный член разбиения Π, так как оппозиция (*завбда*, *завбду*)$_i$ не гомогенна оппозиции (*заводам*, *заводов*)$_i$. Поэтому нельзя определить морфемы по отношению к $\mathfrak{E}_j$. Мы можем, однако, определить квазиморфемы по отношению к $\mathfrak{E}_j$. Существуют квазиморфемы-основания, например *завод-*, *завода-*, *заводо-*, *заводам-*, и дифференциальные квазиморфемы, например *-а*, *-у*, *-ом*, *-е*, *-ам*, *-ы*, *-ах*, *-ами*, *-ми*, *-и*, *-в*, *-х*, *-м*.

Квазиморфема-основание по отношению к $\mathfrak{E}_j$, не содержащая других квазиморфем-оснований по отношению к $\mathfrak{E}_j$, будет называться несводимой квазиморфемой-основанием или, короче, несводимой основой. В приведенном выше примере *завод-* есть несводимая основа.

Квазиморфему-основание, не являющуюся несводимой, будем называть сводимой квазиморфемой-основанием или, короче, сводимой основой. В приведенном выше примере *завода-*, *заводо-* и *заводам-* — это сводимые основы.

Дифференциальные квазиморфемы по отношению к $\mathfrak{E}_j$ могут классифицироваться так: несводимые дифференциальные квазиморфемы, не содержащие других дифференциальных квазиморфем по отношению к $\mathfrak{E}_j$, и сводимые дифференциальные морфемы. В приведенном примере квазиморфемы *—а*, *-у*, *-е*, *-м*, *-ы*, *-х*, *-и*, *-в* несводимы, а квазиморфемы *-ом*, *-ах* и *-ами* сводимы.

Легко проверить следующие положения:

Если $\mathfrak{E}_m$ и $\mathfrak{E}_n$ — это два гомологичных множества, полученных при разбиении Π, то

а) всякая дифференциальная квазиморфема по отношению к $\mathfrak{E}_n$ есть также дифференциальная квазиморфема по отношению к $\mathfrak{E}_m$ и обратно;

б) несводимая дифференциальная квазиморфема по отношению к $\mathfrak{E}_m$ также несводима по отношению к $\mathfrak{E}_n$ и обратно;

в) число квазиморфем-оснований по отношению к $\mathfrak{E}_m$ равно числу квазиморфем-оснований по отношению к $\mathfrak{E}_n$.

Пусть x и y две квазиморфемы по отношению к некоторому множеству $\mathfrak{E}_j$, полученному при разбиении Π. Мы будем говорить, что x и y совместимы, если существует слово из множества $\mathfrak{E}_j$, содержащее цепочку xy. В приведенном выше примере *заводам-* и *-и* совместимы, тогда как *завод-* и *-и* несовместимы.

Ясно, что может существовать совместимость x с y, но не существовать совместимости y с x.

Будем говорить, что x доминирует над y слева в отношении $\mathfrak{E}_j$, и будем писать $(\mathfrak{E}_j)x \rightarrow y$, если квазиморфема x совместима с y в отношении $\mathfrak{E}_j$ и если в каждом слове из множества $\mathfrak{E}_j$, содержащем подцепочку xy, замена xy квазиморфемой x ведет к созданию слова, принадлежащего к множеству $\mathfrak{E}_n$. Например, рассмотрим совокупность $\mathfrak{E}_j$ всех флективных форм слова *завод*. Получим: $(\mathfrak{E}_j)$ *завод* → *-а*, $(\mathfrak{E}_j)$ *завод* → *-у*, $(\mathfrak{E}_j)$ *завод* → *-е*, $(\mathfrak{E}_j)$ *завод* → *-ы*, $(\mathfrak{E}_j)$ *завод* → *-ов*, $(\mathfrak{E}_j)$ *-ам* → *-и*, $(\mathfrak{E}_j)$ *-а* → *-х*, $(\mathfrak{E}_j)$ *завода* → *-ми* и т. д.

Легко видеть, что два гомологичных множества $\mathfrak{E}_m$ и $\mathfrak{E}_n$ изоморфны в отношении доминации слева. Это значит, что между множеством квазиморфем по отношению к $\mathfrak{E}_n$ и множеством квазиморфем по отношению к $\mathfrak{E}_n$ мы можем установить взаимно-однозначное соответствие, сохраняющее отношение доминации слева. Например, если $\mathfrak{E}_m$ есть совокупность флективных форм слова *завод*, а $\mathfrak{E}_n$ есть совокупность флективных форм слова *народ*, то, как мы знаем, $\mathfrak{E}_m$ и $\mathfrak{E}_n$ гомологичны и существует изоморфизм между $\mathfrak{E}_m$ и $\mathfrak{E}_n$ по отношению к доминации слева: цепочка *завод* соответствует цепочке *народ*; *заводами* соответствует *народами*; имеем $(\mathfrak{E}_m)$ *завод* → *-ами*, $(\mathfrak{E}_n)$ *народ* → *-ами* и т. д.

Пусть $\mathfrak{L} = \{\mathfrak{E}, \Pi, \mathfrak{F}\}$ есть некоторый язык, пусть $\lambda \in \mathfrak{E}$, $\mu \in \mathfrak{E}$. Мы будем говорить, что λ и μ находятся в контрастном распределении, если существует фраза из множества $\mathfrak{F}$, содержащая слово λ, в которой замена слова λ словом μ создает фразу, не принадлежащую множеству $\mathfrak{F}$, и существует фраза множества $\mathfrak{F}$, содержащая слово μ, такая, что при замене μ словом λ получится фраза, не принадлежащая $\mathfrak{F}$.

Мы будем говорить, что у языка $\mathfrak{L}$ квазипарадигматическая морфология, если в некоторых двух множествах $\mathfrak{E}_m$ и $\mathfrak{E}_n$, полученных разбиением Π, и для некоторых $\lambda \in \mathfrak{E}_m$, $\mu \in \mathfrak{E}_n$ λ и μ находятся в контрастном распределении.

9. Заключительные замечания

Существует некоторое соответствие между морфемами-основами и квазиморфемами-основами, с одной стороны, семантемами Вандриеса [2] и плеремами Ельмслева [36], с другой стороны. Но наше понимание морфемы и квазиморфемы есть, в первую очередь, разъяснение наиболее распространенных концепций касательно морфемы, концепций, восходящих к И. А. Бодуэну де Куртенэ и с некоторыми изменениями воспринятой и дескриптивной лингвистикой. Речь идет о том понимании, согласно которому морфема есть минимальный звуковой сегмент, обладающий лексическим или грамматическим значением. Критическое рассмотрение этого определения и попытка разъяснить то, что нам казалось имплицитно содержащимся в этом понятии, — все это привело нас к тем определениям, которые даны в настоящей работе. Относительно конкретно-лингвистической стороны дела см. работу [47]. Интересно было бы сопоставить это понятие морфемы с понятиями морфа, фонемы (см., например, [11] и [44]), а также с понятиями и методами, описанными в трудах [25], [30], [33], [35], [4], [37], [39], [7], [9], [43], [10], [31], [29] и [16].

Понятие отношения доминации, употребляемое в настоящей работе, использовалось и в книге С. К. Шаумяна [22]. Существует также простое соотношение между этим отношением доминации и отношением подчинения, рассмотренным в работе [6]. Кроме того, следует сказать, что так называемые «stems», весьма многочисленные в работе [46] Люсьена Теньера, являются иллюстрацией — хотя и не эксплицитной — отношений доминации в синтаксисе.

Представляет интерес для обсуждения и вопрос о расхождениях между моделями парадигмы и квазипарадигмы, предложенными в нашей работе, и парадигмами естественных языков. Так, наша модель не принимает во внимание возможности чередования гласных или согласных, столь частого во многих естественных языках и весьма затрудняющего разделение слова на составные части. См., например, [32] и [34].

Что касается гомологичных классов, они составляют весьма полезный инструмент для прикладной лингвистики (см., например, [41], [7] и [9]) и для структурного языкознания (см., например, [28]).

Очень важен требующий дальнейших исследований вопрос о такой классификации морфем и квазиморфем, которая позволила бы получить формальное соответствие окончания, суффиксов, префиксов и инфиксов. Интересные идеи по вопросу о таком исследовании содержатся в работе [42].

В заключение я хотел бы привлечь внимание к изоморфизму, существующему между понятием п а р а д и г м ы, определенным в настоящей работе, и понятием а б с т р а к т н о г о з в у к а в смысле С. К. Шаумяна ([18], [21]). К понятию абстрактного звука мы можем прийти следующим образом. Рассмотрим множество $\mathfrak{S}$ индивидуальных звуков (в смысле [18]). Во множестве $\mathfrak{S}$ определено отношение эквивалентности R так, что два индивидуальных звука находятся в отношении R, если и только если они отличаются лишь по их позиции. Если будем называть индивидуальным звуком множество физических признаков, то, если даны четыре R-эквивалентных индивидуальных звука s_1, s_2, s_3 и s_4, то оппозиции (s_1/s_2) и (s_3/s_4) гомогенны (т. е. у них есть то же основание). Действительно, общее основание состоит из множества физических признаков, определяющих абстрактный звук, т. е. из множества физических признаков, общих R-эквивалентным индивидуальным звукам; это есть множество физических признаков R-эквивалентных индивидуальных звуков, полученных в отвлечении

от позиции этих звуков. Мы можем, следовательно, сказать, что абстрактный звук есть регулярное множество в том смысле, который этот термин имеет в настоящей работе. Но из того факта, что два различных R-эквивалентных индивидуальных звука, будучи таковыми, несовместимы в одной и той же позиции, мы заключаем, что два R-эквивалентных индивидуальных звука находятся в дополнительном распределении. Значит, абстрактный звук и парадигма допускают ту же логическую модель.

Изоморфизм, о котором мы сейчас говорим, находится в соответствии с выдвинутым С. К. Шаумяном принципом, согласно которому структурное описание должно быть применимо ко всем языковым уровням [21].

ЛИТЕРАТУРА

1. И. А. Бодуэн де Куртенэ. Введение в языковедение. Пг., 1917, стр. 197—198.
2. Ж. Вандриес. Язык. Лингвистическое введение в историю. М., 1937.
3. Г. Глисон. Введение в дескриптивную лингвистику. М., 1959.
4. Л. Ельмслев. Пролегомены к теории языка. «Новое в лингвистике», I, М., 1960.
5. Б. М. Кедров. О числе отношений множеств (понятий). «Логические исследования», М., 1959.
6. О. С. Кулагина. Об одном способе определения грамматических понятий на базе теории множеств. «Проблемы кибернетики», вып. 1. М., 1958.
7. О. С. Кулагина. Об использовании машины при составлении алгоритмов анализа текста. «Проблемы кибернетики», вып. 7. М., 1962.
8. С. Маркус. Теория графов, лингвистические оппозиции и инвариантная структура. «Проблемы структурной лингвистики», М., 1962.
9. И. А. Мельчук. Морфологический анализ при машинном переводе (преимущественно на материале русского языка). «Проблемы кибернетики», вып. 6. М., 1961.
10. И. И. Ревзин. О логической форме лингвистических определений (на примере определения морфемы). «Применение логики в науке и технике» М., 1960.
11. И. И. Ревзин. Модели языка. М., 1962, гл. 3.
12. Ф. де Соссюр. Курс общей лингвистики. М., 1933.
13. Н. С. Трубецкой. Основы фонологии. М., 1960.
14. В. А. Успенский. К определению падежа по А. Н. Колмогорову. «Бюллетень Объединения по проблемам машинного перевода», № 5, 1957.
15. В. А. Успенский. К определению части речи в теоретико-множественной системе языка. «Бюллетень Объединения по проблемам машинного перевода», № 5, 1957.
16. С. Я. Фитиалов. 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.
17. Н. Хомский. Синтаксические структуры. «Новое в лингвистике», II, М., 1962.
18. С. К. Шаумян. Логический анализ понятия фонемы. «Логические исследования», М., 1959.
19. С. К. Шаумян. Двухступенчатая теория фонемы и дифференциальных элементов. ВЯ, 1960, № 5.
20. С. К. Шаумян. Операционные определения и их применение в фонологии. «Применение логики в науке и технике», М., 1960.
21. С. К. Шаумян. Проблемы теоретической фонологии. М., 1962.
22. С. К. Шаумян. Трансформационная грамматика и теория классов слов (в печати).
23. J. Apostel, B. Mandelbrot, A. Morf. Logique, langage, théorie de l'information. Paris, 1957.
24. G. Birkhoff. L attice theory (revised edition). New York, 1948, chapt, 1, § 4.
25. D. L. Bolinger. On defining the morpheme. «Word», vol. 4, 1948.
26. J. Cantineau. Les oppositions significatives. «Cahiers Ferdinand de Saussure», t. 10, 1952, p. 11.
27. J. Cantineau. Le classement logique des oppositions. «Word» vol. 11, 1955.
28. P. Diaconescu. Un mod de descriere a flexiunii nominale, cu aplicaţie la limba romînă contemporană. «Studii şi cercetări lingvistice», vol. XII, nr. 2, 1961.
29. P. Diaconescu. Pe marginea unor lucrări despre morfem. «Studii şi cercetări lingvistice», vol. XIII, nr. 4, 1962.
30. J. Greenberg. Essays in linguistics. New York, 1957.

31. Z. Harris. Morpheme alternants in linguistic analysis. «Language», vol. 18, 1942.
32. Z. Harris. Discontinuos morphemes. «Language», vol. 21, 1945.
33. Z. Harris. From phoneme to morpheme. «Language», vol. 22, 1946.
34. Z. Harris. Methods in structural linguistics. Chicago, 1951.
35. Z. Harris. Distributional structure. «Linguistics to-day», New York, 1954.
36. L. Hjelmslev. Principes de grammaire générale. Kobenhavn, 1928, p. 98 sq.
37. L. Hjelmslev. Essai d'une théorie des morphèmes. «Actes du IV Congrès international des linguistes», Copenhague, 1938.
38. L. Hjelmeslev. La notion de rection. «Travaux du Cercle linguistique de Copenhague», vol. 12.
39. Ch. Hockett. Problems of morphemic analysis. «Language», vol. 23, 1947.
40. A. Martinet. Rôle de la corrélation dans la phonologie dyachronique, «Travaux du Cercle linguistique de Prague», t. VIII.
41. Gr. C. Moisil. Probleme puse de traducerea automată. Conjugarea verbelor în limba romînă scrisă. «Studii şi cercetări lingvistice», nr. 1, 1960.
42. W. Motsch. Zur stellung der Wortbildung in einem formalen Sprachmodell, Thesen über die theoretischen Grundlagen einer wissenschaftliche Grammatik. Berlin, 1961.
43. E. Nida. Morphology: the descriptive analysis of words. Ann Arbor, 1946.
44. S. Saporta. Morph, morpheme, archimorpheme. «Word», vol. 12, 1956.
45. W. Sierpiński. Algébre des ensembles. Warszawa—Wrocław, 1951, p. 136—138.
46. L. Tesnière. Éléments de syntaxe structurale. Paris, 1959.
47. S. Marcus. Modele matematice în lingvistică. Bucureşti, 1963.

THE LOGICAL ASPECT OF LINGUISTIC OPPOSITIONS

Summary

The basic notions of the theory of linguistic oppositions elaborated by Ferdinand de Saussure, Trubetzkoy and Cantineau and the various types of distribution used in descriptive linguistics are expressed in terms of set theory. A classification of opositions from the point of view of their logical structure is given (Table 1), various relations which may occur between oppositions are studied and a number of invariance properties of some types of oppositions by the considered relations are established (Table 2; the «+» indicates invariance, while the «—» indicates non-invariance). It is to be noted that the relations of interdependence, of determination and of constelation used by L. Hjelmslev, may too be expressed by means of the notions which are studied here.

In the second part, the concept of ordered opposition is introduced and the basic notions of morphemic analysis are studied. For further discussions and concrete linguistic illustrations see [47], chapters I and III.

www.ruslang.ru

Б. В. Сухотин

АЛГОРИТМЫ ЛИНГВИСТИЧЕСКОЙ ДЕШИФРОВКИ

ОБЩАЯ ПРОБЛЕМА ДЕШИФРОВКИ. ПОНЯТИЕ НОРМАЛЬНОГО ЛИНГВИСТИЧЕСКОГО АЛГОРИТМА

Дешифровка в том виде, в каком она сейчас существует, мало походит на дисциплину, в какой-то мере научную; наука должна опираться на некоторые общие законы и находить общие методы для решения типичных задач.

Однако имеющиеся успехи в области дешифровки конкретных языков и письменностей обязаны совершенно различным индивидуальным обстоятельствам: находка билингвы — в расшифровке египетских иероглифов; находка словаря — в расшифровке тангутской письменности и наличие словаря при расшифровке письменности майя; удачная догадка Б. Грозного о том, что хеттское слово *ŭatar* соответствует английскому *water*, или, наконец, сходство линеарного письма *B* и кипрского письма вместе с удачной попыткой истолковать тексты как греческие — при расшифровке крито-микенских надписей.

Рассматривая задачу дешифровки вообще, мы не можем рассчитывать ни на какие благоприятные обстоятельства. Задачу дешифровки вообще можно поставить так: имеется некоторое сообщение на каком-то языке. Требуется найти способ перевода его на некоторый известный язык.

В данной фразе мы употребили слово «язык» не в общепринятом смысле. Мы не дадим определения языка, хотя оно представляется нам весьма нужным, но на примерах покажем, что мы имеем в виду.

Все человеческие языки, с нашей точки зрения, также являются языками по отношению ко всей области человеческой активности.

Изображения являются языком по отношению к явлениям видимого мира.

Граммофонная запись есть язык по отношению к звукам.

Предметы, являющиеся результатом производственной деятельности человека, образуют язык по отношению к человеческим потребностям.

Химия есть язык по отношению к превращениям вещества.

Поэтому примером дешифровки может служить и открытие химиком (перевод на язык химии) некоторого химического явления.

Мы считаем, что возможность отыскания способа перевода с одного языка на другой основывается на том, что в строении всех языков имеется существенная общность. Мало того, мы лишь постольку можем один текст считать переводом другого, поскольку строение или взаимосвязь некоторых элементов одного текста (сообщения) повторяют строение или взаимосвязь элементов другого текста.

Поэтому есть надежда найти универсальные законы строения языков. Самые простые и важные из этих законов составляют универсальную грамматику.

С помощью универсальной грамматики можно обнаружить, как выражается то или иное грамматическое явление в конкретном языке, т. е. конкретную грамматику.

Конкретная грамматика не определяет всех видов строения и взаимосвязей элементов данного языка. Известная свобода в строении и взаимосвязях элементов данного языка используется для моделирования конкретной грамматики другого языка средствами данного. Моделируемая грамматика является содержанием данного языка. В частности, некоторый человеческий язык имеет своим содержанием грамматику передаваемой им области действительности.

Моделируемую грамматику тоже можно обнаружить с помощью универсальной, но предварительно надо обнаружить и учесть моделирующую.

Следовательно, с помощью универсальной грамматики можно обнаружить как конкретную грамматику данного текста, так и содержание.

Поэтому общая задача дешифровки совпадает с обнаружением универсальной грамматики.

Под правилом универсальной грамматики мы понимаем алгоритм — точную и общепонятную инструкцию, позволяющую обнаружить данное грамматическое явление в любом тексте на любом языке.

Это традиционное понимание алгоритма уточнено в математике несколькими способами. Мы будем пользоваться тем уточнением, которое предложено А. А. Марковым.

Алгоритмы перерабатывают конечные цепочки (слова), состоящие из некоторых объектов (букв) некоторого конечного набора различных букв (алфавита), в другие слова.

Уточнение понятия алгоритма сводится в основном к тому, что разрешается пользоваться лишь простейшими преобразованиями слов, а именно подстановками некоторой буквы на место вхождения ранее стоявшей буквы.

Алгоритм в уточненном виде А. А. Марков называет нормальным алгорифмом. Понятие нормального алгорифма, очевидно, правильно выражает интуитивное понимание алгоритма (алгорифма), поскольку все доселе известные алгоритмы могут быть нормализуемы. Это значит, что может быть построен нормальный алгоритм, эквивалентный данному. Алгоритмы считаются эквивалентными, если данные слова они перерабатывают одинаковым образом.

Однако это весьма абстрактное понимание алгоритма нас не удовлетворяет. Попытаемся объяснить, почему оно нас не удовлетворяет. Вспомним, что при постановке общей задачи дешифровки мы отказались пользоваться какими-либо сведениями помимо данного сообщения (текста). Допустим, у нас есть текст на человеческом, например, русском языке. Границы между словами не указаны. Пусть требуется составить частотный словарь по этому тексту.

Если прилагается словарь словоформ, то можно без особого труда создать алгоритм и, соответственно, нормальный алгоритм для решения этой задачи.

Но такой алгоритм не удовлетворяет условиям постановки задачи общей дешифровки. Ниже мы дадим определение, направленное на то, чтобы одним алгоритмом могло перерабатываться возможно большее число текстов. Следовательно, в текстах они должны перерабатывать лишь нечто общее, независимое, в первую очередь, от языка. Сам текст должен представляться в весьма простом виде, а алгоритмы — не содержать частей типа словарей словоформ и т. п.

Введем вначале понятие элементарного события и элементарного места.

Пусть у нас имеется некоторый автомат, познающее устройство, (например, человек) различающий n свойств (признаков), каждое из которых может иметь различную интенсивность. Эти признаки предполагаются независимыми и могут служить координатами в n-мерном пространстве.

Если m признаков из n могут характеризовать также и состояние, в котором находится познающее устройство, причем последнее может характеризоваться всеми значениями этих m признаков, то данные m признаков мы назовем аргументами, или m-мерным векторным аргументом, а остальные $n-m$ — вектор-функцией.

Так, например, если в число n признаков входит цвет, а познающее устройство, способное различать цвет, снабжено светофильтром, то цвет входит в число m аргументов. Если же познающее устройство не имеет светофильтров, то цвет входит в число функций. Точку в m-мерном пространстве назовем элементарным местом. Нулевое значение вектор-функции, а также значения, отличающиеся от нулевого на бесконечно малое приращение по одной из $n-m$ координат, назовем элементарным событием. Ввиду того, что познающее устройство имеет конечную разрешающую способность, элементарное место надо трактовать не как точку, а как m-мерный куб с длиной граней, соответствующей разрешающей способности данного автомата по данному признаку. Аналогично надо рассматривать конечные приращения вектор-функции, а не бесконечно малые. Эти конечные величины должны быть минимальными и могут считаться единичными.

Приведем примеры. На телевизионном одноцветном экране различаются высота, ширина и плотность. Первые два признака можно считать аргументом, плотность — функцией, в данном случае скалярной; минимальные размеры квадратика связаны с размером экрана (не меньше $^1/_{625}$ по высоте), минимальное различимое приращение плотности связано с тем, что в телевизоре различаются 64 градации плотности. Элементарным событием в данном случае является белый квадратик соответствующих размеров и слабо-серый квадратик тех же размеров.

Несколько элементарных событий могут занимать одно место. В вышеприведенном примере это соответствует получению всех градаций плотности или умножению элементарного события на число.

В трехцветной печати или в цветной фотографии элементарными событиями являются квадратики трех основных цветов минимальной плотности, а также белый квадратик.

В черно-белом тексте элементарными событиями являются белый и черный квадратик минимальных размеров.

Итак, m-мерное пространство аргументов мы разбиваем на элементарные места — m-мерные единичные кубы. На каждом элементарном месте число всех элементарных событий равно линейной комбинации $a_1x_1+a_2x_2+\ldots+a_rx_r$, где x_i означает один из $n-m$ признаков, а a_i — число элементарных событий по данному признаку.

Нетрудно построить развертку этого множества в одномерную цепочку, состоящую из элементарных событий, поскольку имеется ограниченное подпространство нашего пространства (текст) и число элементарных мест, а также и элементарных событий, занимающих одно место, конечно. Тогда части этой развертки можно считать словами в смысле А. А. Маркова, а сами элементарные события — буквами в том же смысле.

Теперь мы можем дать определение алгоритма, согласующегося с задачей общей дешифровки. Такой алгоритм мы назовем нормальным лингвистическим. Определение будет индуктивным.

1. Нормальным лингвистическим алгоритмом является такой нормальный алгоритм, который перерабатывает слова, состоящие из букв, являющихся элементарными событиями.

2. Нормальным лингвистическим алгоритмом является также нормальный алгоритм, перерабатывающий слова, полученные в результате работы другого нормального лингвистического алгоритма.

При работе над каким-нибудь текстом на человеческом языке использование нормальных лингвистических алгоритмов накладывает следующие требования: помимо текста мы пользуемся только умением различать черное и белое (черную и белую точку). Все остальные сведения можно получать только путем переработки развертки текста в линию, заполненную этими двумя буквами, черной и белой точкой.

Мы предполагаем, что нормальными лингвистическими алгоритмами может быть получено весьма большое количество сведений, имеющих лингвистическое значение.

Определив понятие устойчивого сочетания черных и белых точек, можно определить буквы в обычном лингвистическом смысле как максимально устойчивые в некоторых условиях сочетания черных и белых точек. Так мы обнаружим алфавит данного языка.

Алфавит можно разбивать на классы — например, на гласные и согласные и т. д.

Затем можно определить морфемы как устойчивые сочетания букв, где буквы уже являются буквами в обычном смысле. Мы получим право пользоваться этими буквами, поскольку они будут получены нормальным лингвистическим алгоритмом.

Получив морфемы, можно будет научиться получать слова и, далее, предложения.

При работе по составлению дешифровочных алгоритмов не обязательно, конечно, педантично выдерживать подобную последовательность, но нельзя и слишком грубо нарушать ее, поскольку тогда не будет исключена возможность принятия в качестве данных сведений, которые не могут быть получены нормальным лингвистическим алгоритмом. Нужно помнить, что нарушая подобную последовательность, т. е. забегая вперед, мы выдаем вексель на восполнение недостающей цепочки алгоритмов.

Алгоритмы, предлагаемые читателю в настоящей статье, а также в статье «Экспериментальное выделение классов букв с помощью электронной вычислительной машины» («Проблемы структурной лингвистики», М., 1962), не являются в буквальном смысле слова нормальными лингвистическими алгоритмами, но легко превращаются в таковые. Так, алгоритмы классификации букв и выделения морфем становятся нормальными лингвистическими алгоритмами при создании всего одного алгоритма: алгоритма выделения букв в лингвистическом смысле. Этот алгоритм нами в данный момент разрабатывается. Алгоритм установления связей станет нормальным лингвистическим после значительно большего количества алгоритмов: алгоритма выделения букв, имеющегося алгоритма выделения морфем, по-видимому, алгоритма классификации морфем на лексические и грамматические, алгоритма выделения основ, алгоритма выделения предложений и, наконец, алгоритма классификации слов.

Однако мы верим, что подобные алгоритмы могут быть созданы. Почти все предлагаемые алгоритмы построены по одному принципу. Определяется множество допустимых решений, ответов на задачу, таких, что какой-нибудь из них может оказаться правильным. Затем определяется на этом множестве функция выгодности — такая числовая функция, которая, по нашему предположению, принимает минимальное или максимальное значение в том случае, когда найдено правильное допустимое решение (ответ задачи). Затем определяется процедура, позволяющая

найти решение, для которого функция выгодности минимальна (максимальна). Тем самым находится ответ.

Необходимо указать, что можно построить много алгоритмов для решения одной и той же задачи и качество их может быть весьма различным. Качество алгоритмов определяется только путем проведения экспериментов на известных языках, путем сопоставления с интуитивно полученными решениями. Считается, что интуитивная точность весьма высока и, по-видимому, машинные эксперименты будут конкурировать с человеческими решениями (на известных языках!!) лишь тогда, когда память машин и быстродействие при решении определенных задач будут приближаться к аналогичным свойствам человека.

Составитель алгоритмов должен, разумеется, стремиться к получению максимально точных ответов, добиваясь при этом простоты вычислений и уменьшения их объема. Эти качества тоже обнаруживаются при экспериментальной проверке.

Поэтому очевидна важность экспериментальной работы на вычислительных машинах. Однако экспериментальная работа несравненно более трудоемка, чем само создание алгоритмов. Поэтому интерес экспериментаторов к этим алгоритмам был бы в высшей степени желательным.

АЛГОРИТМ [ВЫДЕЛЕНИЯ] МОРФЕМ

Попытка построения алгоритма выделения морфем как сочетаний некоторых более простых единиц предпринимается не впервые. Такой алгоритм пытался построить З. Харрис в статье «От фонемы к морфеме». По ряду причин, которые мы не имеем возможности здесь излагать, его попытку следует признать неудачной. Некоторые критические замечания по этой статье можно найти в сборнике «О точных методах изучения языка».

Помимо теоретического значения алгоритм выделения морфем может иметь значение средства дешифровки, ценное в том случае, если дешифруемый текст не имеет пробелов. Естественно поэтому желать, чтобы он работал над буквами, а не над фонемами. Кроме того, важно, чтобы вычисления были настолько негромоздкими, чтобы была возможность осуществить их на современных вычислительных машинах (требование практичности).

Разумеется, окончательное суждение о качестве алгоритма можно получить только после проведения эксперимента по выделению морфем с помощью данного алгоритма из текста на известном языке.

Предлагаемый нами алгоритм основан на следующих положениях: 1) морфемы суть в некотором смысле устойчивые сочетания букв; 2) морфемы не пересекаются, т. е. если в цепочке морфем некоторая буква входит в предшествующую морфему, то она не входит в последующую (обратное имеет место в цепочках «чайнворд»).

Допустимым решением (в смысле второго положения) является любое разбиение текста границами. Наилучшим решением будет являться то из допустимых решений, которое образует максимально устойчивые (в определенном ниже смысле) цепочки букв между границами (морфемы).

Определим устойчивость. Цепочки, встречающиеся в тексте один раз, устойчивыми не считаются. Для цепочек букв, встречающихся более двух раз, устойчивость определяется следующим образом: пусть дана цепочка *АБВГДЕ*. Образуем все возможные разбиения этой цепочки на две одной границей:

А	*БВГДЕ*
АБ	*ВГДЕ*
АБВ	*ГДЕ*
АБВГ	*ДЕ*
АБВГД	*Е*

Если длина цепочки равна n, то таких разбиений будет $n-1$. Подсчитаем частоты цепочек *А*, *БВГДЕ*, *АБ*, *ВГДЕ*, *АБВ*, *ГДЕ*, *АБВГ*, *ДЕ*, *АБВГД*, *Е*. Разделим частоту цепочки *АБВГДЕ* на каждое из этих чисел. Мы получим $2(n-1)$ чисел. Среднее арифметическое из этих чисел и будет устойчивостью данной цепочки.

Раскроем смысл этой величины. Что означает частное от деления частоты цепочки *АБВГДЕ* на частоту цепочки *А*? Оно показывает долю употреблений цепочки *А* в составе цепочки *АБВГДЕ* относительно всех употреблений цепочки *А*, или, при неограниченно увеличивающемся тексте, условную вероятность цепочки *АБВГД* при условии, что появилась цепочка *А*. Иными словами, частное от деления частоты цепочки *АБВГДЕ* на частоту цепочки *А* показывает, насколько появление цепочки *А* прогнозирует появление цепочки *АБВГДЕ*.

Каждая из $2(n-1)$ таких дробей показывает, насколько соответствующая часть (*А*, *АБ*, *АБВ*, *АБВГ*, *АБВГД*, *БВГДЕ*, *ВГДЕ*, *ГДЕ*, *ДЕ*, *Е*) прогнозируют целое (цепочку *АБВГДЕ*). Естественно считать, что в более устойчивом сочетании части сильнее прогнозируют целое, чем в менее устойчивом (части реже встречаются самостоятельно). Для сравнения устойчивости разных цепочек нужна величина, характеризующая в целом прогнозируемость данной цепочки ее частями. В качестве такой величины мы употребляем среднюю арифметическую из всех $2(n-1)$ прогнозируемостей. Возможно, что иные средние дадут лучшие результаты, но это не легко решить до эксперимента (имеются в виду средние геометрическая, квадратическая и др.).

Используя этот критерий, можно построить алгоритм выделения морфем в заданном ограниченном тексте. Вначале мы должны найти все сочетания букв, встречающиеся не реже двух раз. Среди этих сочетаний будет много пересекающихся, т. е. противоречащих второму условию. Возможна, например, такая ситуация: несколько раз повторились цепочки *поле*, *тоже*, *лето*, и в тексте имеется фраза: . . .*поле тоже*. . . Можно ли считать, что в этой фразе имеется морфема, вернее слово, *лето*? Если да, то не верно, что здесь налицо морфемы *поле* и *тоже*. Имеются, таким образом, противоречащие решения. Для выбора одного из них используются данные об устойчивости цепочек таким образом, чтобы в окончательном решении в качестве морфем были выбраны максимально устойчивые цепочки.

Перейдем к систематическому изложению алгоритма. Алгоритм состоит из двух частей. В первой части мы находим все сочетания букв, встречающиеся более одного раза, подсчитываем их частоты и устойчивости. Во второй части мы из списка сочетаний с их частотами и устойчивостями вычеркиваем неправильные (не являющиеся морфемами с точки зрения данного алгоритма) сочетания. Алгоритм позволяет указать, если это имеет место, в каком месте текста данное сочетание правильно, в каком неправильно.

Ч а с т ь I. Составление частотного словаря всех сочетаний, встречающихся не реже двух раз.

На первый взгляд, эта задача кажется почти невыполнимой. В самом деле, частотный словарь естественно составлять следующим образом: дать список единиц (словарь, алфавит) и искать каждое слово списка в тексте.

Так можно поступать при подсчете частот букв алфавита и даже при подсчете частот морфем или слов, если они уже известны, потому что число объектов списка ограничено, хотя иногда и велико. Так же можно составлять и словарь частот двухбуквенных сочетаний; при алфавите в 30 букв алфавит двухбуквенных сочетаний будет насчитывать 900 сочетаний.

При увеличении длины сочетаний их алфавиты будут быстро расти: для трехбуквенных сочетаний букв того же алфавита список состоит

из 27 000 сочетаний, для четырехбуквенных — из 810 000 и так далее.

Поэтому надо отказаться от попыток составлять списки сочетаний заданной длины и искать какой-то обходной путь.

При построении требуемой процедуры мы исходим из следующего: если, например, сочетание из трех букв *АБВ* встретилось в данном тексте пять раз, то двухбуквенные сочетания *АБ* и *БВ*, входящие в его состав, должны встретиться по крайней мере столько же, т. е. пять раз или больше. Иными словами, трехбуквенные сочетания надо искать на тех же местах текста, где и входящие в его состав двухбуквенные сочетания. При этом можно учитывать лишь те двухбуквенные сочетания, которые стоят в начале многобуквенных.

Итак, предлагается следующая процедура: вначале составляется полный список двухбуквенных сочетаний и подсчитываются их частоты. Пусть при этом сочетание *АБ* встретилось 10 раз. Выпишем все эти сочетания в столбик:

АБ
АБ
АБ
АБ
АБ
АБ
АБ
АБ
АБ
АБ

Отыщем затем в тексте первое вхождение сочетания *АБ*. Заметим, какая буква стоит справа от этого первого вхождения *АБ* в тексте. Пусть это будет буква *В*. Припишем эту букву к первому сочетанию *АБ* в столбце,

АБВ
АБ
АБ
АБ
. . .
. . .

Рассмотрим теперь второе вхождение сочетания *АБ* в тексте. Заметим, какая буква стоит справа от этого вхождения. Пусть это будет буква *Я*. Припишем ее ко второму сочетанию в нашем столбце:

АБВ
АБЯ
АБ
. . .

Обратим внимание на эти продолженные сочетания *АБВ* и *АБЯ*. Допустим, что они входят в совпадающие (повторяющиеся) более длинные сочетания. Однако ясно, что в эти более длинные повторяющиеся сочетания они могут входить лишь в качестве последних букв, что-нибудь вроде . . .*РНОАБ*. . . в обоих случаях, ибо следующие буквы в тексте разные (*В* и *Я*).

Перейдем к третьему вхождению сочетания *АБ* в тексте. Заметим, какая буква стоит справа от этого вхождения. Пусть это будет *В*. Припишем ее справа к третьему сочетанию нашего столбца:

АБВ
АБЯ
АБВ
АБ
. . .

Рассмотрим сочетания в столбце. Мы видим, что в нем имеются два совпадающих трехбуквенных сочетания *АБВ*. Это — прямое указание на то, что в тексте повторяются по крайней мере трехбуквенные сочетания, содержащие *АБВ*. Но интересно узнать, в каких максимально длинных повторяющихся сочетаниях содержится *АБ*.

Заметив совпадение продолжений третьего и первого вхождений сочетаний *АБ*, снова продолжим первое вхождение на одну букву:

АБВГ
АБЯ
АБВ
АБ
. . .

и снова третье:

АБВГ
АБЯ
АБВГ
АБ
. . .

Совпали первое и третье продолжение? Да, совпали. Тогда снова вернемся к первому вхождению и продолжим его еще на одну букву:

АБВГД
АБЯ
АБВГ
АБ
. . .

и снова к третьему:

АБВГД
АБЯ
АБВГО
АБ
. . .

На этом этапе мы получили несовпадение продолжений. Это является свидетельством того, что повторяются многобуквенные сочетания, кончающиеся на *АБВГ*, и все входящие в них сочетания.

Так заканчивается процесс продолжения одновременно первого и третьего сочетаний вправо.

Затем надо продолжать первое и четвертое, первое и пятое — до первого и последнего включительно. После этого следует продолжать второе и третье, второе и четвертое — до второго и последнего. Потом продолжаются третье, четвертое и все остальные со всеми последующими. В тех случаях, когда не совпадают уже имеющиеся продолжения, надо переходить к следующим парам.

На основании этого списка можно подсчитать частоты многобуквенных сочетаний, содержащих *АБ*.

Отбросим последнюю букву в сочетаниях данного списка:

АБВГ
АБ
АБВГ
АБ
АБ
АБ
Б
. . .

Теперь нетрудно подсчитать, сколько раз встретились сочетания, содержащие *АБ* на первом месте. Для этого проверяется, сколько раз сочетание, записанное в первой строке, встречается в строках, расположенных ниже; та же процедура повторяется для сочетаний во второй, третьей и остальных строках.

После обработки всего списка для расширений *АБ* вычеркивается повторяющаяся информация.

Необходимо указать на следующую особенность этой процедуры. Как правило, некоторое сочетание с его частотой будет даваться в качестве отдельной словарной статьи только в том случае, если не существует сочетания, включавшего данное и имевшего ту же частоту (в противном случае частота включенного сочетания будет равна частоте включающего сочетания). Однако для некоторых сочетаний эта лишняя информация будет получена, поскольку списки расширений делаются для всех двухбуквенных сочетаний, входящих в данное многобуквенное.

Поэтому после составления части частотного словаря по очередному списку надо проверить, не была ли некоторая информация получена уже ранее, и если да, то лишнюю информацию надо уничтожить для экономии места.

Так мы составим частотный словарь всех сочетаний, встречающихся чаще чем один раз.

Для вычисления устойчивостей этих сочетаний нехватает только знания частот букв. Но если частотный словарь двухбуквенных сочетаний представить в виде квадратной таблицы, то частота каждой буквы получится в результате сложения всех чисел, стоящих в соответствующей строке этой таблицы.

Затем для каждого сочетания можно вычислить устойчивость по способу, предложенному выше.

Ч а с т ь II. Отыскание правильных сочетаний (максимально устойчивых, т. е. морфем с точки зрения данного алгоритма).

Построим этот способ по индукции. Мы считаем правильным сочетание, имеющее самую большую устойчивость. Такое сочетание всегда будет в словаре устойчивостей; в весьма редких случаях оно может быть не единственным.

Пусть таким сочетанием окажется *АБВГД*. Отыщем в тексте все вхождения этого сочетания. Из того, что сочетание *АБВГД* правильно, следует, что ни одно сочетание в данном месте текста, имеющее начало или конец внутри *АБВГД*, неправильно. Поэтому, если в частотном словаре записано сочетание, начинающееся на *ВГД* и имеющее своим продолжением буквы, следующие за данным вхождением *АБВГД*, то частоту этого сочетания надо уменьшить на единицу и должным образом пересчитать устойчивость. Отыскивать сочетания, имеющие начало или конец в данном вхождении сочетания *АБВГД*, следует так: отделим одну букву слева и посмотрим, имеется ли в списке (в частотном словаре) сочетание, начинающееся с $БВГДХ_1$. Пусть такое сочетание имеется — $БВГДХ_1 \ldots X_k$ и оно по длине меньше других сочетаний, начинающихся на $БВГДХ_1$. Частоту этого сочетания уменьшаем на единицу и ищем сочетания, начинающиеся на $БВГДХ_1 \ldots X_k$; если такое сочетание есть, поступаем, как в предыдущем случае, если нет, то ищем сочетания, начинающиеся на $ВГДХ_1$. Так просматриваем все сочетания влево и вправо. Когда частота некоторого сочетания становится равной единице, данное сочетание вычеркивается из частотного словаря. После каждого уменьшения частоты некоторого сочетания на единицу следует пересчитать устойчивости всех тех сочетаний, устойчивость которых зависит от частоты данного.

Так надо просмотреть все вхождения сочетания *АБВГД*, уменьшить частоты сочетаний, пересекающихся с этими вхождениями, одновременно пересчитывая устойчивости. После этого в частотном словаре не будет

отражено ни одно сочетание, пересекающееся с найденным правильным. Он снова, как и до работы, будет удовлетворять второму условию (морфемы, т. е. правильные сочетания не пересекаются).

Из остальных сочетаний частотного словаря снова выберем сочетание с максимальной устойчивостью и снова внесем изменения в частотный словарь, в соответствии с просмотром всех вхождений данного сочетания. Так нужно обработать весь частотный словарь. Сочетания, вычеркнутые в этом процессе, всегда неправильны с точки зрения нашего алгоритма. Сочетания, частота которых уменьшалась, неправильны в тех вхождениях, на основании разрыва которых эта частота уменьшалась. Ввиду того, что на каждом этапе мы выбирали максимально устойчивое сочетание, достигается максимальная устойчивость разбиения текста в целом.

Отметим одну особенность процедуры вычеркивания неправильных сочетаний. Допустим, есть неправильное сочетание *АБВГДЕ* и пересекающееся с ним правильное сочетание *ДЕЖЗИК*. Частота сочетания *АБВГДЕ* (а если записано сочетание *АБВГД*, то и частота *АБВГД*) должна быть уменьшена на единицу. В то же время частоту сочетания *АБВГ* на единицу уменьшать не надо, так как *ДЕЖЗИК* с ним не пересекается. Но может оказаться, что сочетания *АБВГ* нет в частотном словаре, поскольку в качестве словарной статьи заносились лишь такие из частей более длинных сочетаний, которые по частоте отличны от включающих их сочетаний.

Если сочетания *АБВГ* действительно нет в качестве отдельной статьи, то его надо внести в словарь с прежней частотой.

Способ работы с длинными текстами. Основной смысл этого способа заключается в следующем: небольшой фрагмент с начала текста объявляется базисным. Весь остальной текст считается дополнительным. Из дополнительного текста выбираются только те сочетания, которые имеются и в базисном тексте. Благодаря этому сильно снижается число сочетаний, рассматриваемых одновременно. Затем в качестве базисного можно выбрать следующий фрагмент текста, а в качестве дополнительного — то, что за ним следует; таким образом можно просмотреть весь текст. Этот способ входит в первую часть алгоритма — составление частотного словаря.

Вначале подсчитываются частоты двухбуквенных сочетаний для базисного текста. Затем к ним прибавляются частоты двухбуквенных сочетаний дополнительного текста, но только для тех сочетаний, частоты которых не были равны нулю в базисном тексте.

Затем полученные двухбуквенные сочетания расширяются обычным образом, но сочетания из дополнительного текста расширяются только за счет сравнения с вхождениями данного сочетания в базисный текст.

В качестве результата работы нашего алгоритма должны выделяться морфемы, а также сочетания морфем: слова, словосочетания. Кроме того, будут выделяться сложные буквы, вроде нем. *sch*.

В настоящее время разработан алгоритм для различения этих объектов.

Морфемы, состоящие из одной буквы, можно будет выделить, проводя границы, соответствующие выделенным ранее более длинным морфемам.

Нулевые морфемы данным алгоритмом выделяться не могут.

Основная задача по дальнейшему улучшению этого алгоритма состоит в повышении точности, т. е. числа правильных выделений морфем относительно всех выделений морфем; эффективности, т. е. числа правильных выделений морфем относительно числа морфем, имеющихся в данном тексте; практичности, т. е. сокращению времени работы машины по программе, составленной на основе этого алгоритма.

Изложенный алгоритм дает первое приближение к морфемам языка. В настоящее время этот алгоритм усовершенствован. Предполагается, что результаты его работы будут значительно более точными.

АЛГОРИТМ УСТАНОВЛЕНИЯ СВЯЗИ СЛОВ В ПРЕДЛОЖЕНИИ

«С чем связано это слово?» — такой вопрос нам приходилось слышать часто, а особенно часто от преподавателя латинского языка. Ответить на него было не просто. Во-первых, язык малознакомый, во-вторых, не ясно, как это «связано».

В этой главе предлагается процедура установления связи между словами, на основании которой мы сможем более ясно представить, что́ такое связь. Процедура включает операции, достаточно простые для того, чтобы их могла выполнить вычислительная машина, а процесс вычисления длинен. Поэтому осуществление этой процедуры по самой ее сути предназначено машине.

В настоящем алгоритме допущения довольно сильны, но для получения многих требуемых сведений у нас уже имеются алгоритмы.

В настоящем алгоритме предполагается известным деление слов на классы вроде классов Р. Л. Добрушина [1] и границы предложения (всегда имеются в виду простые предложения). При этом не надо знать, что класс 1 глагол, а класс 2 — наречие; достаточно знать, что данное слово принадлежит к классу 1, а не 2.

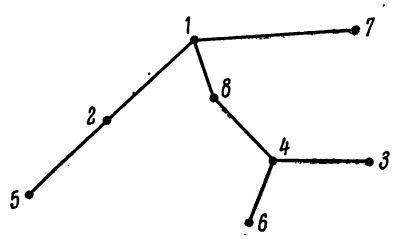
Мы считаем свои допущения приемлемыми, потому что у нас есть, хотя и не вполне совершенный, алгоритм выделения морфем; кроме того, мы верим в то, что алгоритм выделения слов и алгоритм выделения предложений при известных морфемах создать не трудно.

Поясним на примере, каких результатов мы ожидаем от алгоритма обнаружения связей слов. Пусть в нашем тексте встретилось предложение *Большой дом стоял на крутом берегу реки.*

Примем следующие обозначения:

Глаголы	1
Именительный падеж существительного	2
Родительный падеж существительного	3
Предложный падеж существительного	4
Именительный падеж прилагательного	5
Предложный падеж прилагательного	6
Наречия	7
Предлоги	8

Тогда наш пример можно будет записать так: 521786643. Нарисуем схему этого предложения, подобную той, которую нам показывали в школе:



[1] Р. Л. Добрушин. Математическая лингвистика. «Математическое просвещение», 1961, № 6.

Однако классы, получаемые при такой классификации, очень мелки, и для экспериментов мы рекомендуем гораздо более крупные и менее «правильные», например для русского языка: глагол в личной форме, каждый из падежей существительного, каждый из падежей прилагательного или причастия как отдельные классы, инфинитив, деепричастие, наречие, предлог, частица, союз (см. ниже).

Черта между 5 и 2 показывает, что сочетание *большой дом* осмысленно; то, что нет черты между 5 и 8, показывает, что сочетание *большой на* не имеет смысла.

Наш алгоритм должен найти именно такую схему для каждого предложения нашего текста, не пользуясь критерием осмысленности. Разумеется, схемы следует искать для предложений, содержащих более двух слов. В дальнейшем уточненное понятие схемы будет обозначаться словом «граф».

Ввиду того, что этот алгоритм предусматривает знание простых предложений, а в тексте простые предложения могут быть разорваны вводными и придаточными предложениями, при кодировке тексты надо несколько перестраивать. Именно: все простые предложения, входящие в состав данного сложного, следует записывать друг за другом.

Настоящий алгоритм состоит из двух частей. В первой части для всего текста составляется таблица особых величин, нужных для второй части. Во второй части с помощью этих величин разбирается каждое предложение текста. В результате разбора для каждой пары слов, входящих в данное предложение, указывается, связаны эти слова или не связаны.

Итак, для начала работы нам нужен текст, закодированный следующим образом: 1) каждому слову соответствует номер или символ данного класса, 2) между простыми предложениями стоит символ, обозначающий границу.

Составляется квадратная таблица, число строк и столбцов которой равно числу классов. Эта таблица заполняется согласно следующим правилам: i-тый столбец озаглавлен так же, как и i-тая строка; при последовательном обзоре всех предложений текста, если в данном предложении, где найден i-тый символ, имеется и j-тый символ (один или, безразлично, несколько), то к числу, стоящему в i-той строке и j-том столбце, прибавляется единица. До работы все клетки таблицы заполнены нулями. Затем числа, стоящие в каждой строке, складываются, и каждое число, стоящее в i-той строке, делится на сумму, полученную для i-той строки (нормируется по частотности i-того символа).

Числа, полученные таким образом, показывают долю появлений i-того символа, сопровождаемого появлением j-того символа, ко всем проявлениям i-того символа. При неограниченном удлинении текста пределы, к которым стремятся эти числа, суть условные вероятности j-того символа при условии, что i-тый символ появился. Иными словами, каждое такое число показывает, насколько i-тый символ прогнозирует j-ый.

В таблице, разумеется, найдется и число, показывающее, насколько j-ый символ прогнозирует i-тый. Среднюю арифметическую этих двух величин мы будем использовать в качестве характеристики взаимной прогнозируемости i-того и j-того символов независимо от направления.

Изобразим таблицу прогнозируемостей.

	1	2	3	4	. . .
1	$\frac{x_{11}}{\Sigma_1}$	$\frac{x_{12}}{\Sigma_1}$	$\frac{x_{13}}{\Sigma_1}$		. . .
2	$\frac{x_{21}}{\Sigma_2}$	$\frac{x_{22}}{\Sigma_2}$	$\frac{x_{23}}{\Sigma_2}$		. .
3	.	.	.		. . .
4	.	.	.		. . .
.					

Индексы при x читаются «один-один», «один-два», а не «одиннадцать», «двенадцать» и т. д., и означают, что данный x стоит в строке 1 и столбце 1,

в строке 1 и столбце 2 и т. д. Индекс при сигме означает, что сигма относится к строке 1, 2 и т. д.

Буква x означает число, показывающее, сколько раз символ, озаглавливающий строку, и символ, озаглавливающий столбец, встречаются одновременно в предложениях нашего текста. Сигма означает частоту данного символа. Каждая дробь показывает, насколько символ, озаглавливающий строку, прогнозирует символ, озаглавливающий столбец.

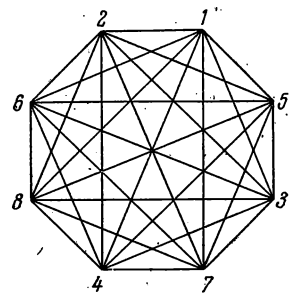
Для того, чтобы получить из этой таблицы таблицу взаимных прогнозируемостей, надо эту таблицу повернуть на 90° и каждое число новой таблицы сложить с соответствующим числом старой таблицы и разделить на 2. Полученная таблица будет симметрична относительно главной диагонали, т. е. число, стоящее в i-той строке и j-том столбце, будет равно числу, стоящему в j-той строке и i-том столбце.

В первоначальной таблице все прогнозируемости заменим взаимными прогнозируемостями. На этом кончается первая часть алгоритма.

Перейдем ко второй части, к разбору отдельных предложений. Для изложения этой части нам понадобятся некоторые понятия теории графов. Эти понятия мы введем описательно, отсылая за уточнениями к статье Г. Н. Поварова [2].

Граф, например сеть авиалиний, марсианских каналов, есть пара множеств, одно из которых («стрелки») находится в одно-двузначном отношении к другому («вершины»). Среди вершин каждой стрелки различают начало и конец. Стрелки, такие, что при некоторой нумерации начало $i+1$ стрелки совпадает с концом i-той стрелки, образуют путь из начала k-той стрелки в конец l-той ($k \leqslant i \leqslant l-1$). Связанными вершинами графа называются такие две вершины, что: 1) от одной из них к другой проходит по крайней мере один путь; 2) каждая вершина считается связанной сама с собой, хотя бы и не было пути, ведущего из этой вершины в нее саму. Число стрелок, составляющих путь, называется длиной пути. Вершины, связанные путем длины 1, называются непосредственно связанными. Граф, все вершины которого связаны, называется связным.

Теперь представим некоторое разбираемое предложение в виде графа. Для этого нарисуем символы, входящие в предложение, в вершинах многоугольника, имеющего столько вершин, сколько слов в предложении (все вхождения одного символа надо перенумеровать индексами и записать как отдельные вершины). Каждую вершину соединим со всеми остальными. Тогда для некоторого предложения из восьми символов мы получим такой рисунок:



Граф на нашем рисунке — связный; кроме того, добавление новых стрелок не породит путей, отличных от уже имеющихся, а будет лишь дублировать старые, т. е. на этом графе представлены все возможные виды путей, соединяющие восемь вершин.

---

[2] Г. Н. Поваров. О матричном анализе связей в частично ориентированных графах. «Успехи математических наук», 1956, т. 11, вып. 5.

В самом деле, предположим, что найдется путь, не имеющийся в нашем графе. Тогда этот путь отличается от любого пути, имеющегося в графе, тем, что, по крайней мере, одна пара вершин a и b, соединененная стрелкой и входящая в этот путь, не соединена стрелкой в графе. Но это противоречит тому, что любую вершину, в том числе a, мы соединяли со всеми остальными, в том числе с b. Случай, когда несходство заключается в том, что непосредственно не связанная в этом пути пара вершин непосредственно связана в графе, исключается, потому что в графе можно выбрать путь, в котором эти вершины также не связаны непосредственно.

Поэтому очевидно, что все вершины, связанные в лингвистическом или в любом смысле, будут связаны и на нашем графе, конечно, при том условии, что это понятие связанности будет частным случаем изложенного выше.

Однако в нашем графе будут и такие стрелки, которые не соответствуют лингвистическим связям. Если мы их выбросим с помощью вполне определенной операции, то мы, во-первых, получим граф лингвистических связей предложения, а во-вторых, сможем определить само понятие лингвистической связи.

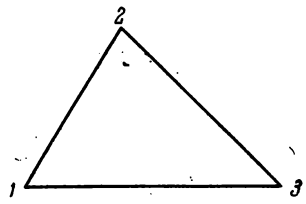
Эту операцию, являющуюся второй частью нашего алгоритма, мы излагаем ниже.

Поставим в соответствие стрелке, соединяющей i-тую и j-тую вершины (индексы не учитываются), взаимную прогнозируемость i и j.

Рассмотрим стрелку, для которой прогнозируемость минимальна. Если выбрасывание этой стрелки не превращает граф в несвязный, то выбрасываем ее и переходим к рассмотрению следующей стрелки (в порядке возрастания прогнозируемости). В противном случае переходим к рассмотрению следующей стрелки без выбрасывания предыдущей. Таким образом просматриваются все стрелки. Легко видеть, что в оставшемся графе средняя из всех взаимных прогнозируемостей пар символов максимальная сравнительно с другими связными графами данного предложения. Такой граф мы называем максимально устойчивым, при этом понимание устойчивости близко к предложенному И. А. Мельчуком. Граф должен оставаться связным, так как естественно считать, что в предложении все слова связаны, хотя не всегда непосредственно.

Теперь мы можем дать определение лингвистической связи. Лингвистическая связь есть отношение на множество слов, входящих в данное предложение. Считается, что слова a и b объединяются этим отношением тогда и только тогда, когда слова a и b непосредственно связаны в максимально устойчивом графе этого предложения.

Отметим одну особенность максимально устойчивых графов. В теории графов циклом называется такой путь, начало и конец которого совпадают. Связный граф, не имеющий циклов, назовем неориентированным деревом. Легко видеть, что графы, получающиеся с помощью нашего алгоритма, суть неориентированные деревья. Допустим, что это не верно, что в графе есть цикл, например:



Но тогда мы можем выбросить одну стрелку, например 13, если взаимная прогнозируемость минимальна для 13, а граф останется связным. Значит, граф не был еще преобразован до конца, не является максимально

устойчивым. Это противоречит условию, и поэтому наше допущение неверно.

В том случае, когда несколько из подлежащих выбрасыванию стрелок соответствуют одной и той же взаимной прогнозируемости, ответ может оказаться многозначным. Это значит, что для одного предложения мы получим несколько максимально устойчивых графов.

Алгоритм в изложенном виде нагляден, но неудобен для работы, так как требует вычерчивания графов и подсчета путей. Это неудобство особенно существенно для программирования на машинах. Однако имеется простая вычислительная методика, которую можно найти в статье Г. Н. Поварова. Она основана на представлении графов матрицами.

Графу из n вершин ставится в соответствие квадратная матрица n-ого порядка. Стрелке, соединяющей i-тую вершину с j-той, ставится в соответствие единица, занесенная в клетку на пересечении i-той строки и j-того столбца. Число путей из i в j, называемое квазиминором по $ij(Q_{ij})$, определяется по особому алгоритму, который мы даем в примечании [3].

Итак, вместо того, чтобы рисовать граф разбираемого предложения, мы будем составлять его матрицу. Исходная матрица для примера приве-

---

[3] В этом примечании мы будем пользоваться обозначениями Г. Н. Поварова. В основном тексте мы их не применяли ввиду некоторой сложности.

Квадратную матрицу порядка p обозначим $\| a_{ij} \|$, где $1 \leqslant i \leqslant p$, $1 \leqslant j \leqslant p$, a_{ij} — элемент матрицы, стоящий в i-той строке и j-том столбце. Обозначим через $\| a_{ij} \nearrow_{lk} \|$ матрицу, в которой ликвидирован l-тый столбец и k-тая строка, но номера строк и столбцов оставлены прежними. Квазиминор по kl обозначим через $a_{ij} \nearrow_{lk} \mid_{kl}$ и определим следующим образом: квазиминор $\mid a_{ij} \nearrow_{lk} \mid_{kl}$ есть сумма произведений вида:

$$\Sigma a_{ki_1} a_{i_1 i_2} \ldots a_{i_r l},$$

где суммирование происходит по всем размещениям $i_1 i_2 \ldots i_r$ $(r = 0.1, \ldots, p-2)$. Если $k = l$, то $\mid a_{ij} \nearrow_{lk} \mid_{kl} = a_{kk}$. Вычислять квазиминоры можно с помощью следующего правила: если $k \neq l$, то $\mid a_{ij} \nearrow_{lk} \mid_{kl} = \sum\limits_{m \neq k} a_{km} A^{(l)}_{km}$, где $A^{(l)}_{km} = 1$ при $m = l$ и $A^{(l)}_{km} = \mid a_{ij} \nearrow_{lk} \nearrow_{km} \mid_{ml}$ при $m \neq l$. Приведем образец вычисления квазиминора $a_{ij} \nearrow_{43} \mid_{34}$, например, для следующей матрицы:

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 2 & 0 & 3 & 1 & 1 \\ 0 & 3 & 0 & 0 & 2 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix}$$

$$\mid a_{ij \to 43} \mid_{34}$$

$$\begin{array}{|c|cccc|}\hline & 1 & 2 & 4 & 5 \\ \hline 1 & 0 & 1 & 1 & 0 \\ 2 & 2 & 0 & 1 & 1 \\ 3 & 0 & 3 & 0 & 2 \\ 5 & 0 & 1 & 0 & 0 \\ \hline \end{array}_{34} = 3 \begin{array}{|c|ccc|}\hline & 1 & 4 & 5 \\ \hline 0 & 1 & 0 \\ 2 & 1 & 1 \\ 0 & 0 & 0 \\ \hline \end{array}_{24} + 2 \begin{array}{|c|cc|}\hline & 1 & 2 & 4 \\ \hline 0 & 1 & 1 \\ 2 & 0 & 1 \\ 0 & 1 & 0 \\ \hline \end{array}_{54} =$$

$$= 3 \cdot 2 \begin{array}{|c|cc|}\hline & 4 & 5 \\ \hline 1 & 1 & 0 \\ 5 & 0 & 0 \\ \hline \end{array}_{14} + 3 + 3 \begin{array}{|c|cc|}\hline & 1 & 4 \\ \hline 1 & 0 & 1 \\ 5 & 0 & 0 \\ \hline \end{array}_{54} + 2 \begin{array}{|c|cc|}\hline & 1 & 4 \\ \hline 1 & 0 & 1 \\ 2 & 2 & 1 \\ \hline \end{array}_{24} = 3 \cdot 2 + 3 + 2 \cdot 2 \begin{array}{|c|c|}\hline & 4 \\ \hline 1 & 1 \\ \hline \end{array}_{14} +$$

$$+ 2 = 15$$

Номера строк и столбцов отчеркнуты.

денного на стр. 85, будет иметь вид:

	1	2	3	4	5	6	7	8
1	0	1	1	1	1	1	1	1
2	1	0	1	1	1	1	1	1
3	1	1	0	1	1	1	1	1
4	1	1	1	0	1	1	1	1
5	1	1	1	1	0	1	1	1
6	1	1	1	1	1	0	1	1
7	1	1	1	1	1	1	0	1
8	1	1	1	1	1	1	1	0

После каждого выбрасывания стрелки, которое теперь выражается в постановке нуля вместо единицы в матрице предложения, подсчитывается $\frac{1}{2}(n^2-n)$ квазиминоров для клеточек выше главной диагонали.

Условие связности графа будет выражено тем, что ни один квазиминор, кроме квазиминоров вида Q_{ii}, не должен быть равен нулю. То, что результирующий граф будет деревом, выразится в том, что для его матрицы все квазиминоры, кроме Q_{ii}, будут равны единице. В связи с тем, что нас интересует лишь равенство — неравенство нулю наших квазиминоров, для их вычислений можно пользоваться логическим сложением и умножением по правилам: $1+1=1+0=0+1=1$, $0+0=0$; $1\cdot 0=0\cdot 1=0\cdot 0=0$, $1\cdot 1=1$.

Стрелкам на графе, полученном таким образом, можно приписать направления. Для этого в таблице прогнозируемостей, еще не приведенной к симметричному виду, следует выбирать максимальное из двух чисел: x-прогнозируемости a по $в$ и y-прогнозируемости $в$ по a. Если $x<y$, то стрелке, соединяющей вершины x и y, приписывается направление от x к y. В самом деле, если b сильно прогнозирует b, «не может обойтись без x», то можно считать, что y зависит от x, а x подчиняет y. Если $y<x$, то имеет место обратное.

В настоящее время разработан алгоритм установления связей, являющийся уточнением данного и учитывающий лексические признаки слов.

Кодировка текстов для алгоритма установления лингвистических связей. Кодировка текстов для алгоритма установления лингвистических связей складывается из двух этапов. Первый этап — перевод слов текста в символы классов слов. Второй этап — преобразование сложных предложений к такому виду, при котором ни одно простое предложение, входящее в его состав, не входило бы в состав другого простого предложения.

Первый этап не может быть пока формализован ввиду того, что не существует алгоритма определения классов слов. Этот этап должен осуществляться кодировщиком. В дальнейшем, по-видимому, классы слов могут быть выделены алгоритмом, аналогичным алгоритму определения классов букв.

Поэтому пока мы можем рекомендовать только наше интуитивное деление на классы. Если экспериментатор предпочтет иное деление на классы, то в существе дела ничего не изменится; однако наше деление мы старались выбрать с тем, чтобы результаты работы алгоритма были бы достаточно интересными, но не было бы излишних трудностей при программировании и кодировке.

Приводим список классов для русского языка:

1. Глагол в личной форме
2. Именительный
3. Родительный
4. Дательный
5. Творительный
6. Предложный

(2–6: падежи существительного или местоимения)

7. Именительный
8. Родительный
9. Дательный
10. Творительный
11. Предложный

(7–11: падежи прилагательного или причастия или местоимения)

12. Деепричастие
13. Инфинитив
14. Наречие или соответствующее местоимение
15. Предлог
16. Союз
17. Частица (кроме *да*, *нет*)

Междометия и частицы *да*, *нет* не кодируются, потому что они образуют однословные предложения.

Рассмотрим второй этап кодировки текстов. При кодировке текстов необходимо отмечать границы предложений символами типа круглых скобок, например: (*человек* (*с которым Иван работал после того*) (*как я кончил школу*) *уехал в Киев*).

То же самое в символах классов: (2(15. 10. 2. 1. 15. 3) (16. 2. 1. 4) 1. 15. 4). Преобразование текстов производится следующим образом: запишем слева исходное предложение, а справа оставим место для преобразованного:

(2(15. 10. 2. 1. 15. 3) (16. 2. 1. 4) 1. 15. 4)

Перепишем справа на оставленном месте цепочки, заключенные между начальной и конечной круглой скобками, если внутри них нет еще скобок, а слева, под исходным предложением, исходное предложение с выброшенными цепочками, которые были записаны справа:

(2.1.15.4) (15.10.2.1.15.3) (16.2.1.4)

Затем преобразованное исходное предложение (оно записано слева) нужно снова преобразовать по тому же правилу:

(15.10.2.1.15.3) (16.2.1.4) (2.1.4.15)

Это преобразование производить до тех пор, пока исходное предложение не окажется пустым. Второй этап можно запрограммировать и осуществить на машине.

КОРРЕКТИРУЮЩИЙ АЛГОРИТМ К АЛГОРИТМУ ВЫДЕЛЕНИЯ КЛАССОВ БУКВ

Описание и машинные эксперименты

Напомним результаты работы первого алгоритма выделения классов букв.

Русский текст. Гласные: *а*, *о*, *у*, *ы*, *е*, *и*, *я*, *ю*, *ь*, *ъ*, *э*; согласные: *б*, *в*, *г*, *д*, *ж*, *з*, *й*, *к*, *л*, *м*, *н*, *п*, *р*, *с*, *т*, *ф*, *х*, *ц*, *ч*, *ш*, *щ*.

Французский текст. Гласные: *а*, *о*, *е*, *i*, *у*, *и*, *к*; согласные: *b*, *c*, *d*, *f*, *g*, *h*, *j*, *l*, *m*, *n*, *p*, *q*, *r*, *s*, *t*, *v*, *w*, *x*, *z*.

Английский текст. Гласные: *а*, *о*, *и*, *i*, *е*, *у*, *t*; согласные: *b*, *c*, *d*, *f*, *g*, *h*, *j*, *k*, *l*, *m*, *p*, *n*, *q*, *r*, *s*, *v*, *w*, *x*, *z*.

Немецкий текст. Гласные: *а*, *о*, *и*, *i*, *е*, *ä*, *ö*, *у*, *s*, *h*, *k*; согласные: *b*, *c*, *d*, *f*, *g*, *j*, *l*, *m*, *n*, *p*, *q*, *r*, *t*, *v*, *w*, *x*, *z*, β, *ü*.

Испанский текст. Гласные: *а*, *о*, *и*, *i*, *е*, *á*, *ó*, *и*, *í*, *é*, *у*; согласные: *b*, *c*, *d*, *f*, *g*, *h*, *j*, *k*, *l*, *m*, *n*, *p*, *q*, *r*, *s*, *t*, *v*, *w*, *x*, *z*, *ñ*.

В результате, того что кодировка производилась людьми, не знающими языков, без должного контроля, во французском тексте в качестве отдельных букв не были закодированы гласные с диакритическими значками; в испанском, напротив, были отдельно закодированы гласные с ударениями; в русском тексте в половине материала были смешаны ь и ъ (перепутан их порядок в алфавите).

Если отвлечься от некоторых несообразностей, вызванных этими ошибками, то результаты можно оценить так: для русского текста результаты безошибочны (ъ попал в гласные вследствие путаницы с ь, который на письме все еще выполняет функцию гласного); для испанского текста — также безошибочны; для французского результаты почти безошибочны (*k*, попавшее в гласные, встретилось всего раз в тексте из 100 000 букв, видимо, в иностранных словах).

Существенные ошибки встретились в результатах для английского и в особенности для немецкого текстов. Ошибки касаются весьма частых букв и объясняются наличием сочетаний букв, передающих один звук: в английском *th*, в немецком *sch*, *ck*, *ch*. Занесение *ü* в согласные объясняется, вероятно, ошибками в квалификации *s*, *h*, *k*.

Для английского языка был проведен также счет на половинном тексте, в 5 000 букв. По сравнению с основным результатом добавилась одна ошибка: *g* было включено в гласные, видимо, вследствие наличия частого сочетания *ng*.

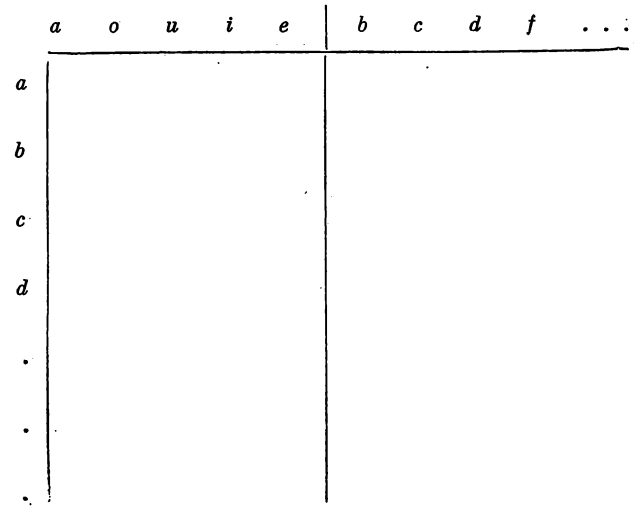
Если ошибки типа неверного занесения *k* во французском несущественны, то для исправления регулярных ошибок желательно создать корректирующий алгоритм.

Очевидно, что корректирующий алгоритм должен использовать иные дистрибутивные признаки, различающие гласные и согласные, по сравнению с основным алгоритмом.

Корректирующий алгоритм, излагаемый ниже, опирается на следующую гипотезу: дана цепочка из пяти букв $x_1x_2x_3x_4x_5$. Если буква x_3 принадлежит к классу 1, то обыкновенно либо равное количество букв окрестности $x_1x_2 \ldots x_4x_5$ принадлежит к классу 1 и к классу 2, либо большинство букв окрестности принадлежит к классу 2.

Случай, когда равное количество (по две буквы) букв окрестности принадлежит к классу 1 и к классу 2, включает в себя предположение первого алгоритма выделения букв на классы. В самом деле, тогда мы предполагали, что за гласными наиболее обычно следуют согласные, значит, цепочка из пяти букв в типичном случае должна была бы иметь вид *ГСГСГ*. В окрестности центральной буквы гласных и согласных поровну. В корректирующем алгоритме, как мы увидим, случай равенства количества гласных и согласных не используется.

Алгоритм работает следующим образом. Строится квадратная таблица, строки и столбцы которой озаглавлены так: строки озаглавлены буквами данного алфавита в произвольном порядке; столбцы озаглавлены буквами данного алфавита в порядке, определяемом разбиением букв на классы, полученным с помощью первого алгоритма, например вначале гласные, потом согласные. Между гласными и согласными проведена граница (т. е. запоминается, начиная с какого столбца идут столбцы, соответствующие гласным), например:



Перед началом работы во все клетки таблицы помещаются нули. Затем с начала текста отсчитывается пять букв. Отыскивается строка, соответствующая третьей букве в этой пятерке. Отыскиваются столбцы, соответствующие остальным буквам этой пятерки. В клетки, стоящие на пересечении найденной строки и столбцов, прибавляется единица к ранее стоявшим числам. Затем отсчитывается пятерка, начиная со второй буквы и т. д.

Разберем, в какие части таблицы попадут эти единицы. Пусть средняя буква в цепочке — гласная и пусть она правильно определена предшествующим алгоритмом. Тогда в силу нашего предположения большее число единиц попадает в половину таблицы с «согласными» столбцами. Если средняя буква в цепочке согласная и если она правильно определена предшествующим алгоритмом, то, в силу нашего предположения, большое число единиц попадет в половину таблицы с «гласными» столбцами.

Пусть теперь средняя буква цепочки — гласная, но предшествующий алгоритм определил ее как согласную. По нашему предположению, болшее чйсло единиц должно быть занесено в половину таблицы с «согласными» столбцами. С другой стороны, если мы поверили предшествующему алгоритму, то мы ожидаем, что большее число единиц будет занесено в половину таблицы с «гласными» столбцами. Увидев, что это не так, мы обнаруживаем ошибку первого алгоритма.

В случае с неправильно определенной согласной ситуация аналогична.

Изложим наш алгоритм еще раз в виде правил.

1. Составить квадратную таблицу по числу букв в алфавите.
2. Озаглавить строки буквами в алфавитном порядке.
3. Озаглавить первые m столбцов буквами, определенными первым алгоритмом как гласные (класс 1).
4. Озаглавить остальные столбцы буквами, определенными как согласные.
5. Взять первые пять букв с начала текста.
6. К числам, стоящим на пересечении строки, озаглавленной средней буквой в этой пятерке со столбцами, озаглавленными остальными буквами этой пятерки, прибавить по единице.
7. Проверить, есть ли в тексте за выделенной пятеркой еще буквы. Если нет, то перейти к пункту 10.
8. Сдвинуть начало текста на одну букву вправо.
9. Перейти к пункту 5.
10. Для каждой строки подсчитать разность сумм чисел, стоящих на пересечении данной строки с «согласными» столбцами, с суммами чисел, стоящими на пересечении данной строки с «гласными» столбцами.

Немецкий

	e	*a*	*i*	*u*	*o*	*h*	*ö*	*ä*	*k*	*s*	*y*	*ü*	*l*	*b*
a	257	45	58	121	17	156	1	1	53	228	1	7	164	67
b	174	67	51	28	22	28		5	4	45		18	39	2
c	117	83	134	55	34	317	5	7	30	148	1	14	61	11
d	488	136	163	128	56	47	4	13	2	99	1	9	47	16
e	570	257	553	152	76	284	15	18	120	558	3	24	251	174
f	96	60	38	44	19	16	8	7	3	19		9	20	1
g	326	109	94	85	21	25	3	7	23	65	1	3	50	26
h	284	156	211	95	66	28	10	19	9	15	1	13	70	28
i	553	58	66	35	24	211		5	55	128		2	128	51
j	25	15	5	12	1	8				14			3	
k	120	53	55	23	55	9	3	7	2	10	1	13	42	4
l	251	164	128	54	57	70	6	12	42	84		10	86	39
m	177	92	86	48	49	27	1	3	19	22	2	5	25	5
n	856	275	290	194	106	1	8	24	63	187	1	23	80	59
o	76	17	24	14	8	66		2	55	65	1		57	22
p	43	21	12	15	13	4	2	7	2	25	2	2	14	4
q	2			1				1						1
r	704	192	186	131	127	115	3	22	48	145	1	35	75	88
s	558	229	228	106	65	150	5	12	10	124	3	18	84	45
t	512	220	190	110	79	124	12	21	38	201		18	102	27
u	152	121	35	10	14	95	2	5	23	106		3	54	28
v	72	8	13	4	28	9			1	7			6	1
w	145	81	76	26	17	38	3	2	1	40		5	26	9
x														
y	3	1			1	1			1	3				
z	115	22	51	64	8	7	3	3	6	9	1	4	13	4
β	23			7	11	4		1	1	12		1	7	1
ö	15	1		2		10			3	5			6	
ä	18	1	5		2	19			7	12			12	5
ü	24	7	2	3	3	13			13	18			10	18

Француз

	e	*a*	*o*	*i*	*u*	*y*	*k*	*h*	*c*	*d*	*f*	*g*	*b*
a	408	87	62	208	162	25	1	31	122	132	42	55	46
b	77	46	46	22	17				2	4	1		
c	268	122	111	78	57	4	1	62	28	32	4	1	2
d	383	132	74	85	121	4	1	6	32	28	6	6	4
e	625	408	213	326	415	12	1	72	268	384	49	64	77
f	49	42	25	43	32			1	4	6	10		
g	64	55	17	24	19				1	6	1		
h	72	31	34	22	22				62	6	1		
i	326	208	160	90	136	8	5	22	78	85	43	24	22
j	36	10	19	16	16				3	6			
k	1	1	1	5	3				1	1			
l	488	241	86	224	177	17	5	19	56	41	19	0	34
m	282	108	100	102	27	12		22	39	25	1	7	5
n	488	285	334	203	196	7		21	124	196	20	31	7
o	213	62	16	160	180	17	1	34	111	74	25	17	46
p	229	163	96	43	88	13		3	13	30		8	
q	94	23	7	50	104			2	5	2		1	
r	680	323	201	201	210	12	2	28	72	117	24		1
s	716	342	289	352	263	12	1	19	86	130	12		9
t	656	299	227	216	162	2		7	68	76	6	7	9
u	415	161	180	136	32	5	3	22	57	121	32	19	17

текст

c	*d*	*f*	*g*	*j*	*t*	*m*	*n*	*p*	*q*	*r*	*v*	*w*	*x*	*z*	*β*
83	136	60	109	15	220	92	274	21		191	8	80		22	20
11	16	1	26		27	5	59	4	1	88	1	9		4	1
6	32	1	7	3	91	14	65			81	5	30		5	2
32	36	13	36	4	74	32	273	3	1	201	8	10		18	12
117	488	96	326	25	512	177	856	43	2	704	72	145		115	23
1	13	32	15		16	8	46	9		35		3		6	
7	36	15	24	4	54	19	170	1		85	3	22		16	2
317	47	16	25	8	124	27	127	4		115	9	38		7	4
139	163	38	94	5	19	86	290	12		186	13	76		51	13
3	4		4		8		9			5					
30	20	3	23		38	19	63	2		48	1	1		6	1
61	47	20	50	3	102	25	80	14		75	6	26		13	7
14	32	8	19		62	32	106	6	1	64	6	4		16	5
66	273	46	170	9	190	106	238	19		147	22	68		5	3
34	56	19	21	1	79	49	106	13		127	28	17		8	11
	3	9	1		10	6	19	4		34				4	
	1					1				1					
81	201	35	85	5	201	64	147	34	1			0 1		41	11
148	99	19	65	14	201	22	187	25		145		40		19	12
91	74	16	54	8	154	62	190	10		201	8	22		64	4
55	128	44	85	12	110	48	194	15	1	131	4	26		64	7
5	8		3		8	6	22			73	2			3	1
30	10	3	22		22	4	69			101				16	2
1	1		1			2	1	2		1				1	
5	18	6	16		64	16	50	4		41	3	16			1
2	12		2		4	5	3			1	1	2		1	
5	4	8	3		12	1	8	2		13				3	
7	13	7	7		21	3	24	7	1	22		2		3	1
14	9	9	3		18	5	23	2		35		5		4	1

ский текст

j	*l*	*m*	*n*	*p*	*q*	*r*	*s*	*t*	*v*	*w*	*x*	*z*
10	241	108	285	163	23	323	342	299	92		27	5
	34	5	7			61	29	9	1		2	1
3	56	39	124	13	5	72	85	68	22		5	2
6	41	25	196	30	2	116	130	76	29		8	3
36	488	282	488	229	94	680	716	656	143		51	15
	19	1	20			24	12	6	2		2	
	20	7	31	8	1	32	17	7	1		1	
	19	22	21	3	2	28	19	7	2		2	1
16	224	102	203	43	50	201	252	216	73		20	1
	3	4	5	2		5	10	8	7			2
	5					2	1				3	
3	96	43	123	72	15	39	226	102	27		7	4
4	43	80	101	22	1	49	46	53	1		2	
5	123	101	132	53	21	117	237	249	65		2	3
19	36	100	334	96	7	201	289	227	66		5	2
2	72	22	53	26	2	138	119	42	3		9	2
	15	1	21	2		19	22	10			2	
5	99	49	117	138	19	150	199	251	51		1	2
10	226	46	237	119	22	199	228	159	33		9	1
8	102	53	249	42	10	25	159	106	23		2	
16	177	27	186	88	104	210	263	162	54		35	

Француз

	e	a	o	i	u	y	k	h	c	d	f	g	b
v	143	92	66	73	55	3		2	22	29	2	1	1
w													
x	51	16	5	20	35	2	3	2	5	8	2	1	2
y	12	25	17	3	5				4	4			
z	15	5	2	1		1			1	2	3		1

Английский

	e	o	a	i	t	u	y	g	b	k	c	f
a	160	60	57	64	138	14	35	54	24	13	30	29
b	54	22	24	11	23	27	12	10	2	1	6	5
c	57	41	30	15	12	22	7	2	6	8	2	5
d	136	91	121	73	37	39	14	17	15	5	9	5
e	196	144	160	140	224	17	48	49	54	26	57	42
f	42	67	29	32	36	8	9	7	5		5	2
g	45	61	54	71	29	15	2	8	10	1	2	7
h	248	86	72	123	212	32	32	44	12	7	39	24
i	140	30	64	12	121	18	15	71	11	28	25	32
j					1	3						
k	26	42	13	18	11	2	1	1	1		8	
l	106	123	92	83	40	21	38	16	12	17	17	28
m	65	46	53	39	26	19	4	4	5	1	11	6
n	130	124	150	157	117	46	22	74	16	21	17	13
o	144	100	59	30	151	65	22	61	22	42	41	67
p	48	31	53	44	10	12	6	3		2	1	3
q	1	1	1	4		4	2					
r	205	104	87	82	77	32	27	22	19	11	14	29
s	161	110	150	127	90	45	20	27	14	9	26	30
t	224	151	138	121	84	75	21	29	23	11	12	36
u	27	65	14	18	75	4	9	15	27	2	22	3
v	67	11	23	27		1	4	2	1		1	1
w	88	69	78	58	36	3	23	9	1	2	8	11
x	2											
y	48	22	35	15	21	9	2	2	12	1	7	9
z			1									

Английский

	e	o	a	i	y	u	t	b	c	d	f	g
a	347	108	51	109	71	25	292	44	67	253	51	104
b	100	42	44	24	20	46	38	6	9	34	12	19
c	104	74	67	57	8	32	19	9	2	14	11	4
d	298	193	253	165	33	69	96	30	14	64	22	32
e	478	280	347	243	107	55	501	100	104	298	93	95
f	93	133	51	73	13	18	61	12	11	22	16	10
g	95	103	104	135	8	45	59	19	4	32	10	10
h	534	177	167	248	64	65	449	23	74	153	47	83
i	243	66	109	30	21	37	241	24	57	165	73	135
j	1					5	1	1	1	2		
k	55	87	29	46	3	2	21	2	12	14	3	1
l	227	223	195	145	78	40	99	17	31	89	57	28
m	140	89	110	66	18	37	47	12	21	35	9	18
n	303	232	301	304	45	90	237	28	32	207	26	156
o	280	190	107	66	63	135	298	42	74	193	133	103
p	126	73	98	83	10	21	34	2	6	15	11	4
q	3	7	2	13	3	13				1		1

ский текст

j	*l*	*m*	*n*	*p*	*q*	*r*	*s*	*t*	*v*	*w*	*x*	*z*
7	27	1	65	3		51	33	23			2	3
	7	2	2	9	2	1	9	2	2			
	17	12	7	13		12	12	2	3		2	1
2	4		3	2		2	1		3			

текст (половина)

h	*j*	*d*	*l*	*m*	*n*	*p*	*q*	*r*	*s*	*v*	*w*	*x*	*z*
72		121	92	53	150	53	1	17	150	23	78		1
12		15	12	5	16			19	10	1	1		
39		9	17	11	17	1		14	26	1	8		
65	1	40	44	14	107	7		35	36	7	14		
241		136	106	65	130	48	1	205	161	67	81	2	
24		5	28	6	13	3		19	30	1	11		
44		17	16	4	75	3		22	27	2	1		
30		65	39	30	36	14		63	126	10	57		1
123		73	83	39	157	42	4	82	127	17	58		
		1			1				1				
7		5	17	1	21	2		11	10		12		
39		44	74	20	50	21	2	23	39	14	32		1
23		14	20	2	24	4		15	41	2	9		
36	1	107	50	24	10	16		38	61	16	45		
86		91	123	46	124	31	1	104	110	11	65		
14		7	21	14	16	16		52	16	19	17	1	
			2								1		
63		35	23	15	38	52		42	62	32	48	1	
126	2	36	39	41	61	16		62	48	6	66		
212	1	37	40	26	117	10		77	90		37		
32	3	39	21	19	46	12	4	31	45	1	8		
10		7	14	2	16	9		32	6		2		
57		14	32	9	45	7	1	98	66	2	4		
						1		1					
32		14	38	4	22	6	2	27	20	4	13	1	
1			1										

текст (полный)

h	*j*	*k*	*l*	*m*	*n*	*p*	*q*	*r*	*s*	*v*	*w*	*x*	*z*
167		29	195	110	301	98	2	181	269	41	137	1	2
23	1	2	17	12	28	2		36	26	1	2		
74	1	12	31	21	32	6		29		2	9	2	
153	2	14	89	35	207	15	1	79	67	17	31	1	
534	1	55	227	140	303	126	3	446	333	125	161	12	1
47		3	57	9	26	11		58	44	1	19		
83		1	28	18	156	4	1	45	45	4	15		
82		10	65	72	177	32	2	130	237	22	110		
248		46	145	66	304	83	13	171	240	42	104		1
					2			19	3				
10			26	2	45	4		45	17		6		
65		26	146	41	85	42	7	34	92	23	56	2	1
72		2	41	6	63	8		72	69	4	15		
177	2	45	85	63	86	21	1	202	112	28	98		1
177		87	223	89	232	73	7	91	195	16	126	1	
32		4	42	8	21	34		1	34	13	13	5	
2			7		1				1		1		

Английский

	e	*o*	*a*	*t*	*y*	*u*	*t*	*b*	*c*	*d*	*f*	*g*
r	446	202	181	172	49	80	156	36	29	79	58	45
s	333	195	269	240	49	84	187	26	40	67	44	45
t	501	298	292	241	46	120	156	38	19	96	61	59
u	55	135	25	37	19	4	120	46	32	69	18	45
v	125	16	41	42	7	1	3	1	2	17	1	4
w	161	126	136	105	33	13	62	2	9	31	19	15
x	12	1	1		1		1		2	1		
y	107	63	71	21	2	19	46	20	8	33	13	8
z	1		2	1	1							

11. Проверить, для всех ли гласных разность больше нуля. Те гласные, для которых эта разность меньше нуля, запомнить. На самом деле, т. е. с точки зрения корректирующего алгоритма, они суть согласные.

12. Проверить, для всех ли согласных разность меньше нуля. Те согласные, для которых разность больше нуля, суть на самом деле гласные, запомнить их.

Вполне возможно было формулировать этот алгоритм как основной, а первый — как корректирующий. Для этого надо всего лишь поменять местами таблицы, входящие в состав обоих алгоритмов.

Эксперименты проводились на четырех текстах: немецком, английском, французском (по 10 000 букв каждый текст) и английском длиной в 5000 букв.

Результаты таковы: в немецком тексте все ошибки были исправлены, т. е. было опознано ошибочное занесение *s*, *h* и *k* в гласные, а *ü* — в согласные; в английском тексте было исправлено ошибочное занесение *t* в гласные, но внесена одна ошибка: буква *j*, весьма редкая, была занесена в гласные. Эта ошибка, видимо, случайная, так как в половинном английском тексте такой ошибки не появилось. В половинном английском тексте были устранены все ошибки, т. е. было опознано ошибочное занесение *t* и *g* в гласные. Во французском тексте не было опознано ошибочное занесение *k* в гласные; кроме того, в гласные была занесена буква *h*.

Таким образом, корректирующий алгоритм выдает ответы или вообще без ошибок, или с ошибками, которые, как правило, менее грубы, чем те, что были ранее.

Хотя эти результаты мы оцениваем как сравнительно хорошие, мы считали желательным создание более точного алгоритма. Мы предполагали, что более точным явится алгоритм, излагаемый ниже. Оказалось, однако, что он эквивалентен — в обоих своих вариантах — алгоритму, опубликованному в сборнике «Проблемы структурной лингвистики», 1962.

Тем не менее, мы считаем целесообразным опубликование этого алгоритма, ввиду того, что эквивалентность обоих алгоритмов не очевидна. Доказательство эквивалентности, а также ряд новых алгоритмов для разделения гласных и согласных будут публиковаться впоследствии.

На стр. 94—98 приведены таблицы для английского, французского и немецкого текстов.

НОВЫЙ АЛГОРИТМ ВЫДЕЛЕНИЯ КЛАССОВ БУКВ

Данный алгоритм предназначен для разделения букв на два класса. Построен он следующим образом: определяется множество допустимых решений, затем из этих решений выбирается наилучшее.

Допустимые решения. Однократное применение алгоритма должно разбить данный алфавит на два класса. Мы принимаем в качестве допустимых решений только разбиение на непересекающиеся классы,

текст (полный)

h	*j*	*k*	*l*	*m*	*n*	*p*	*q*	*r*	*s*	*v*	*w*	*x*	*z*
13		19	45	34	72	91	1	80	125	56	84	2	
237	3	17	92	69	112	34	1	125	94	15	114		
449	1	21	99	47	237	34		156	187	3	62	1	
65	5	2	40	37	90	21	13	60	84	1	13		
22			23	4	28	13		56	15		3		
110		6	56	15	98	13	1	84	114	3	8		
			2			5		2					
64		3	78	18	45	10	3	49	49	7	33	1	1
1			1		1								

т. е. если некоторая буква входит в класс 1, то она не входит в класс 2 и наоборот. Приведем примеры разбиений русского алфавита, являющихся допустимыми решениями:

Класс 1	Класс 2
абвгджз	*ийклмнопрстуфхцчшщьыъэюя*
а	*бвгдежзийклмнопрстуфхчшщьыъэюя*
(пустой)	(весь алфавит)
аоуеиэыьюя	*бвгджзйклмнпрстфхцчшщъ*

В реальных текстах бывают случаи, когда одна буква входит в разные классы. Например, буква *v* в латинских текстах обозначала как согласный *u*, так и гласный *u*. Для обнаружения и учета таких случаев необходимы специальные алгоритмы, которыми мы пока не располагаем.

Наилучшее решение. Наилучшим решением мы считаем такое, которое более всего соответствует интуитивному делению данных лингвистических единиц на классы. Для букв таким делением является деление на буквы, соответствующие гласным и согласным звукам. Для русских букв наилучшим делением (решением) является последнее из приведенных выше.

Допустимые разбиения текста. Назовем сегментом любой отрезок текста, содержащий целое число рассматриваемых лингвистических единиц, в данном случае букв. Назовем сегментами класса 1 сегменты, состоящие только из букв класса 1, сегментами класса 2 — сегменты, состоящие только из букв класса 2. Такие сегменты класса 1, приписывание к которым справа или слева хотя бы одной буквы превращает их просто в сегменты, называются максимальными сегментами класса 1. Аналогично определяются максимальные сегменты класса 2.

Каждое допустимое решение порождает одно и только одно разбиение текста на максимальные сегменты классов 1 и 2; причем за каждым максимальным сегментом класса 1 следует максимальный сегмент класса 2 и наоборот.

Например, первое из вышеприведенных допустимых разбиений порождает следующее разбиение текста стихотворения А. С. Пушкина «Бесы». Цифры под строкой означают классы.

мч а тсятучи в ьютсятучине в и д имкоюлуна . . .
2 1 2 1 2 1 2 1 2

и т. д.

Основные гипотезы. Мы предлагаем несколько гипотез, таких, что принятие одной из них даст соответствующий вариант данного алгоритма. Окончательный выбор гипотезы, очевидно, можно будет сделать лишь после проведения экспериментов.

По первой гипотезе мы предполагаем, что наилучшему решению соответствует разбиение текста, обладающее следующим свойством: разность средней длины максимальных сегментов класса 1 и средней длины максимальных сегментов класса 2 минимальна по абсолютной величине. Наш алгоритм будет излагаться, исходя из этой гипотезы.

По второй гипотезе мы предполагаем, что наилучшему решению соответствует минимальное значение средней длины максимальных сегментов.

Можно предложить весьма большое количество правдоподобных гипотез, например, заменяя один вид средней другим и строя из этих средних величины разной сложности.

Правдоподобность вышеприведенных гипотез очевидна из нижеследующего примера:

м ч а т с я т у ч и в ь ю т с я т у ч и н е в и д и м к о ю л у н а . . .
2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1

Очевидно, что в этом примере максимальные сегменты обоих классов равномернее по длине, чем в предыдущем случае.

Если мы примем эту основную гипотезу, то процедура отыскания наилучшего решения сведется к поиску такого решения, для которого в порожденном им разбиении текста разность средних длин максимальных сегментов обоих классов минимальна по абсолютной величине.

Процедура отыскания наилучшего решения. Считаем первоначально весь алфавит занесенным в класс 2. Разбиение текста, соответствующее этому допустимому решению, содержит один максимальный сегмент, равный по длине всему тексту и принадлежащий классу 2. Можно считать, что налицо и сегмент (сегменты) класса 2, но средняя длина их равна 0. Разность средних длин, таким образом, максимальна и равна длине всего текста. Затем построим такие допустимые решения: в класс 1 включим одну из букв нашего алфавита, а остальные включим в класс 2. Таких решений будет столько, сколько букв в данном алфавите. Для каждого решения построим соответствующее разбиение текста и вычислим разность средних длин максимальных сегментов.

Из этих решений выберем такое, которое максимально уменьшает разность средних длин максимальных сегментов. Это решение мы будем улучшать в процессе дальнейшей работы.

Пусть, например, выбранное решение такое: класс 1—*о*; класс 2 — все остальное.

Построим ряд новых решений так: перенесем из класса 2 в класс 1 еще одну букву. Мы получим такие решения:

Класс 1	Класс 2
оа	(все остальное)
об	(все остальное)
ов	(все остальное)
. . . . .	

Из этих решений выберем такое, которое максимально уменьшает разность, полученную для предыдущего решения. Пусть это будет решение $K1$: *оа*, $K2$: все остальное.

Ряд новых решений получим аналогично, перенеся из класса 2 в класс 1 одну букву. Дальше повторяются все перечисленные выше операции.

Процесс заканчивается тогда, когда ни одно из ряда новых построенных решений не уменьшает разности средних длин максимальных сегментов.

Этот алгоритм является более общим, чем предыдущие, потому, что он не требует, чтобы сочетания единиц разных классов были чаще сочетаний единиц одного класса; достаточно, чтобы группы единиц обоих классов были по возможности равномерны по длине. Поэтому можно надеяться,

что он даст существенно более точные результаты и будет годен для работ в области морфологии и синтаксиса.

В заключение нам хотелось бы выразить надежду на то, что эта работа будет содействовать возникновению интереса к общей задаче дешифровки и выразить искреннюю благодарность тем, кто, помогая автору, не пожалел своего времени и труда: Ю. Д. Апресяну, К. И. Бабицкому, А. М. Великанову, Т. М. Великановой и С. К. Шаумяну.

ALGORITHMS FOR LINGUISTIC DECIPHERING

Summary

It is useful to consider linguistics as a sciense of deciphering for obtaining such definitions of language phenomena as are not based on meaning.

A printed text or a manuscript may be reduced to combinations of coloured and colourless dots, and any other text — to a certain inventory of micro-events of a similar nature.

We may assume such micro-events to be given in advance. Any other information contained in the text should be derived by means of deciphering algorithms (their definition is given in the paper).

The algorithms presented in this paper are not deciphering algorithms in the strict sense of our difinition, but they will obtain this quality after filling in certain gaps.

The algorithms are as follows: 1) an algorithm for obtaining morphemes, 2) an algorithm for finding out the graph of a sentence, 3) two algorithms for classifying letters into classes similar to vowels and consonants.

All those algorithms may be used for an automatic analysis of text by a computor. The results of a machine experiment staged after the third algorithm are also given in the paper.

Ю. Д. Апресян

СОВРЕМЕННЫЕ МЕТОДЫ ИЗУЧЕНИЯ ЗНАЧЕНИЙ И НЕКОТОРЫЕ ПРОБЛЕМЫ СТРУКТУРНОЙ ЛИНГВИСТИКИ

В последние 10—15 лет в лингвистике оживленно обсуждаются различные вопросы, связанные с изучением значений [43], [88], [102], [113], [133], [134], [179]. Во многих написанных в эти годы работах выдвинуты новые важные идеи и методологические принципы. Все большее распространение получают математические и экспериментальные методы изучения значений. Первые скромные достижения в этой области не могут не заинтересовать специалистов, занимающихся структурной лингвистикой и другими разделами современного языкознания. Представляется поэтому своевременным систематически рассмотреть современные принципы изучения значений и уже полученные результаты, сопоставив их с теориями традиционной семантики и выяснив их связь с проблематикой структурной лингвистики. Именно эти цели и преследует автор данного обзора.

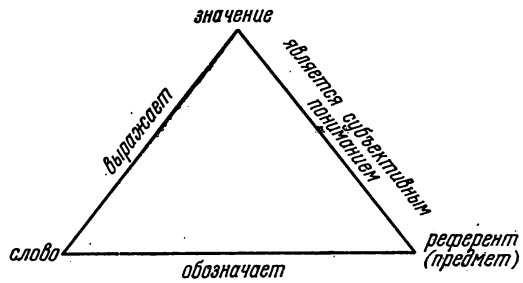
В обзоре учитывается в основном литература последнего десятилетия. Исключение сделано лишь для нескольких классических исследований 30-х и 40-х годов. Особое внимание уделяется тем работам, в которых предлагаются методы, практически применявшиеся при изучении значений в естественных языках.

Далеко не во всех упоминаемых в обзоре работах решаются собственно лингвистические задачи. Естественно поэтому, что мы не можем дать им исчерпывающую оценку. Они интересуют нас с чисто лингвистической точки зрения и рассматриваются лишь в той мере, в какой они связаны с общеязыковедческой проблематикой.

А. ПОСТАНОВКА ВОПРОСА О ЗНАЧЕНИИ

I. Традиционное учение о значении

1. Система Г. Стерна. Классическую законченность традиционному учению о значении придал шведский лингвист Г. Стерн [165]. Исходя из наиболее совершенной в то время семиотической теории К. Огдена и И. Ричардса [138], он построил следующий «треугольник отнесенности», в терминах которого и было определено понятие значения:



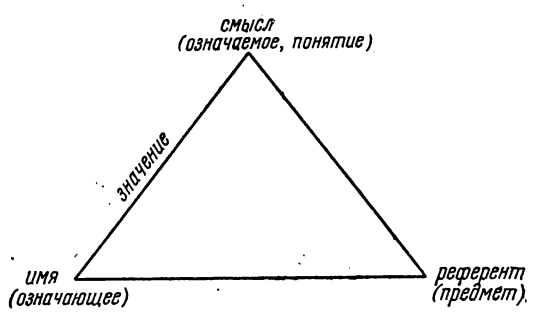
(1) «Значение» помещено здесь в ту вершину треугольника, которую у К. Огдена и И. Ричардса занимает «понятие». Это предопределяет принципиальное отождествление значения с «понятием» аристотелевской логики (их различает лишь то, что в «значение» включаются также некоторые эмоциональные элементы, поскольку оно рассматривается как «субъективное понимание референта»). Значение рассматривается, следовательно, не как отношение, а как некоторая интеллектуальная сущность (вещь), «носителем» которой является слово (т. е. имя). Эта прямолинейная упрощенная концепция значения была усвоена многими современными авторами, даже такими тонкими, как А. И. Смирницкий [50].

(2) Г. Стерн анализирует значение знаменательного слова, а не лингвистической формы вообще. Предметом семантики оказываются, таким образом, «лексические значения»; значения грамматических морфем выводятся за пределы семантики даже в тех случаях, когда означаемыми морфем являются факты действительности (ср. значение множественного числа в русском языке). Это привело, между прочим, к тому, что семантика стала рассматриваться как часть лексикологии. В современных пособиях по лексикологии семантика излагается как один из ее разделов наряду с фразеологией, заимствованиями и словообразованием [1], [6], [9], [15], [34], [51].

(3) Хотя «контекст» используется для установления значения слова, он не является объектом анализа. Объектом анализа является значение отдельно взятого слова, искусственно изолированного из цепочки элементов (фразы), так как считается, что любое слово и вне контекста обладает всем множеством значений, которые оно может принимать в контекстах. Это характерно и для позднейших работ в области традиционной семасиологии, даже таких, в которых употребляется структурная терминология — [63], [115], [171—174]; данный принцип лежит и в основе лексикографической обработки слова.

(4) Хотя сам Г. Стерн рассматривал значение слова в связи с синонимическим рядом, в который оно входит, существо его теоретической концепции, как показывает и «треугольник отнесенности», не запрещает в принципе рассматривать значение изолированно, «в самом себе и для самого себя». Во многих работах по семасиологии оно именно таким образом и рассматривается; именно этот подход освящен почти всей практикой лексикографии (исключение составляют лишь так называемые идеологические словари). В результате анализируется значение элемента, искусственно изолированного из системы элементов (например, семантического поля).

2. Система С. Ульмана. Сознательную попытку преодолеть некоторые из указанных недостатков традиционной семасиологии и примирить ее с принципами структурной и статистической лингвистики сделали С. Ульман [171—174], П. Гиро [95], Г. Кронассер [115], К. Бальдингер [63], А. Рудскогер [157]. В концепциях этих авторов довольно много общего (ср. синтез их идей в [63]); с наибольшей полнотой сущность нового этапа в развитии семасиологии выражена в работах С. Ульмана. В основе данного им определения значения лежит «треугольник отнесенности» К. Огдена и И. Ричардса, модифицированный следующим образом:



Значение определяется как отношение имени к смыслу (эти два термина точно соответствуют соссюровским терминам «означающее» и «означаемое»).

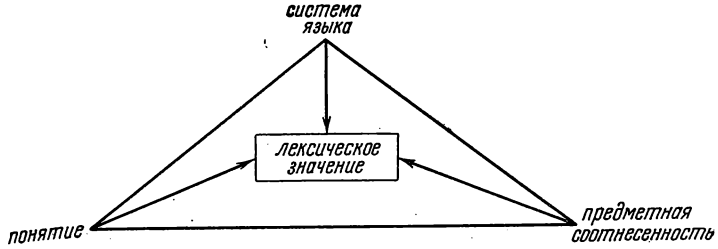
Поскольку значение определено как отношение имени к смыслу, преодолевается отождествление значения и понятия; поскольку при этом рассматриваются всякие имена (означающие), преодолевается переоценка роли лексического значения. Отметим в этой связи, что С. Ульман выделяет формальную и семантическую сторону и в лексикологии и в синтаксисе и предусматривает «синтаксическую семантику» наряду с «лексической семантикой».

Однако С. Ульману не удается преодолеть других недостатков традиционной семантики, несмотря на предпринятые для этого усилия. С. Ульман исходит из того, что имена и смыслы входят в ассоциативные ряды. Ряд, основанный на ассоциации смыслов, он рассматривает вслед за Й. Триром [170] как «семантическое поле». Следовательно, элементом семантического поля является не «функционально» определенное значение, но смысл (понятие), а при анализе значения как такового сохраняет силу «изолирующий» подход традиционной семасиологии. Таким образом, попытка С. Ульмана преодолеть, по его выражению, «кризис семантики» и показать совместимость традиционной семасиологии с новыми структурными методологическими идеями, в частности идеей семантического поля, окончилась неудачей.

3. Учение об «обусловленности» значения. В последние годы была предпринята еще одна попытка модернизировать традиционную семасиологию на основе соссюровской идеи об обусловленности значения лингвистическим рядом (системой знаков). Под обусловленностью понимаются лексические и синтаксические факторы, от которых зависит возможность реализации (лексического) значения того или иного слова. Так, значение «заросший» у прилагательного *глухой* реализуется в сочетании со словами определенной лексической группы — *лес*, *тайга* и т. п., ср. *глухой лес*, *глухая тайга*. Значение «дурак» у слова *осел* нормально требует для своей реализации употребления слова в предикативной функции, ср. *Ну и осел же ты!* Значение «ухаживать» у глагола *бегать* требует для своей реализации предлога *за* с творительным падежом, ср. *он бегал за ней*.

В принципе такие факты были известны давно; они, в частности, весьма тщательно описывались в толковых словарях. Но традиционная семасиология в течение долгого времени проходила мимо этих существенных достижений лексикографической практики. Факты лексической и синтаксической обусловленности значений были проанализированы лишь относительно недавно Дж. Фиртом [85], [86], В. В. Виноградовым [10], Р. С. Гинзбург [16] и некоторыми другими исследователями. Однако никто из них не дал систематического изложения всего комплекса идей традиционной семасиологии в свете принципа обусловленности.

Такую задачу поставил перед собой В. А. Звегинцев [24], исходивший из того, что лексическое значение слова определяется (обусловливается) предметным, понятийным и лингвистическим рядом. Он предложил следующий способ схематизации изложенного представления:



На первый взгляд, в этой схеме отражены все стороны изучаемого объекта, в том числе и те (обусловленность значения лингвистическим ря-

дом, т. е. лексическая и синтаксическая обусловленность), которые не учитывались Г. Стерном, С. Ульманом и их последователями. Фактически, однако, налицо диффузное, нерасчлененное понимание объекта, которое делает невозможным его систематическое исследование. Мы уже не говорим о том, что В. А. Звегинцев возродил преодоленное С. Ульманом представление о лексическом значении как единственном объекте семантики. Таким образом, в концепции В. А. Звегинцева недостатки традиционной семасиологии не только не преодолеваются, но в некоторых отношениях усугубляются.

4. Вопрос о методе. До сих пор мы говорили о некоторых теоретических представлениях, развитых традиционной семасиологией. Если добавить к этому, что в ней отсутствует какое-либо представление о методе и технике исследования, без которых немыслима никакая теория, станет ясно, насколько неудовлетворительным было положение дел в области изучения лингвистических значений по сравнению с изучением лингвистических форм.

Естественно поэтому, что современная лингвистика пошла не по пути адаптации традиционных семантических представлений, но по пути их коренной переработки на основе идей, развиваемых в современной семиотической и логической литературе, причем большое внимание уделяется развитию и освоению новых методов исследования. Эта переработка начата в относительно недавнее время. Поэтому ни о какой стройной и законченной теории лингвистической семантики говорить не приходится, тем более, что разные ученые и школы по-разному подходят к решению основных семантических проблем. Однако уже предложенные теоретические принципы и методологические идеи и разработанные на конкретном материале технические приемы исследования значений заключают в себе так много интересного и поучительного, что представляется полезным их детальный анализ.

II. Современная постановка вопроса о значении

(1) В современной семиотике значение лингвистической формы не отождествляется ни с понятием, ни с предметом, ни с какой-либо другой сущностью, но рассматривается как отношение (подробнее см. пункт 5).

(2) Считается, что значением в этом смысле обладает всякая лингвистическая форма, т. е. лексическая или грамматическая морфема, синтаксический показатель, конструкция, предложение и т. п. По поводу последнего заметим, что уже Г. Фреге рассматривал предложения в качестве сложных имен, имеющих значение и смысл [7]. Таким образом, объектом анализа являются и грамматические и лексические значения.

(3) Значение элемента рассматривается как определенное место в синтагматической цепи элементов. Л. Витгенштейн писал по этому поводу: «Только предложение имеет смысл; только в контексте предложения имя обладает значением» [13]. В связи с этим в современной лингвистике и логике подчеркивается роль употребления как семантической единицы, реально присутствующей в тексте, и значение целого предложения как фактор, определяющий значения составляющих его элементов [23], [67], [83], [85], [109], [155], [178]. Г. Рейхенбах и У. Куайн идут еще дальше, утверждая, что достаточно очевиден только смысл высказывания и что поэтому можно сопоставлять лишь значения целых высказываний, но не значения их частей.

(4) Значение элемента рассматривается как определенное место в системе элементов (семантическом поле, семантическом пространстве). Каждый элемент входит, как правило, сразу в несколько семантических полей; поскольку каждое семантическое поле характеризуется определенным семантическим свойством («семой»), значение элемента может быть

представлено как сумма или пучок элементарных значений (сем, семантических компонентов, семантических множителей). Предполагается, таким образом, что значения образуют структуру. Справедливость этой гипотезы была показана для грамматики и некоторых областей лексики (цветообозначений, терминов родства и т. п.). Впрочем, существуют такие методы трактовки значений, которые позволяют рассматривать как структуру совокупность всех значений (ср. особенно «метод тезаурусов»).

(5) Считается, что значение — сложный объект, скрывающий несколько более простых объектов. Этот сложный объект должен быть надлежащим образом расчленен, чтобы быть изученным. Исследования Р. Карнапа, У. Куайна, Ч. Морриса, Ч. Осгуда, Б. Рассела, А. Черча и других авторов в области семиотики и логической семантики [7], [17], [29], [57], [130], [147], [148—151], [158] позволяют говорить о следующих видах значений [1]: а) Структурное значение — отношение знаков к другим знакам. Отношение знаков к другим знакам на синтагматической оси может быть названо синтаксическим значением или валентностью; отношение знаков к другим знакам на парагматической оси может быть названо дифференциальным значением или значимостью (в полном соответствии с идеей Ф. де Соссюра). б) Сигнификативное значение—отношение знаков к сигнификатам; сигнификат — термин Ч. Морриса; он является экспликатом традиционного термина «содержание понятия» и соответствует термину «смысл» у Г. Фреге и А. Черча, термину «интенсионал» у Р. Карнапа, термину «значение» у У. Куайна. в) Денотативное значение — отношение знака к денотату (вещи), который он обозначает. Термин «денотат» принадлежит А. Черчу; он является экспликатом традиционного термина «объем понятия» и соответствует термину «значение» у Г. Фреге, термину «экстенсионал» у Р. Карнапа, термину «референция» у У. Куайна [2].

Эти три вида значений характеризуют знак независимо от того, является ли он языковой реакцией на внешний стимул или языковым стимулом, вызывающим некоторое внеязыковое поведение, т. е. независимо от того, занимает ли он положение r или s в блумфильдовской цепочке $S \rightarrow r \rightarrow s \rightarrow \rightarrow R$ [70]. Учет положения знака в этой цепочке дает другую семиотическую классификацию значения: а) ситуативное значение — языковая реакция (знак), вызываемая внеязыковым стимулом (денотатом). Этот вид значения в точности соответствует денотативному значению и поэтому объединяется с ним. б) Прагматическое значение — внеязыковая реакция (поведение), вызываемая языковым стимулом (знаком). В него включаются также оценочные и эмоционально-экспрессивные компоненты значения. Данный вид значения не является собственно лингвистическим [17], но изучается некоторыми школами современной психолингвистики и поэтому рассматривается в настоящей работе [3].

Таким образом, современная лингвистика и ряд смежных с ней дисциплин занимаются следующими видами значений: структурными (синтаксическими и дифференциальными), сигнификативными, денотативными и прагматическими. По вопросу о том, каким образом соотносятся описанные множества значений, в современной литературе нет определенной точки

---

[1] Излагаемая классификация является определенной модификацией нескольких существующих классификаций. Используемые в ней термины употребительны, но не общеприняты.

[2] У. Куайн, впрочем, пользуется и карнаповскими терминами «интенсионал» и «экстенсионал» [148—150].

[3] Возможна и другая систематизация видов значений. Интересная сводка существующих классификаций дана в сборнике «Психолингвистика» [145]. В недавнее время оригинальную и глубокую концепцию лингвистического значения предложил И. А. Мельчук [39], [41]. Некоторые ее детали совпадают с изложенными выше представлениями (в частности, предложенное им понятие «синтаксического значения» близко понятию «валентности», а понятие «лексического значения» охватывает понятия «сигнификативного» и «денотативного» значения).

зрения, и мы не будем здесь его обсуждать. Некоторые соображения по этому поводу излагаются нами в разделе Е.

Всякая лингвистическая форма обладает по крайней мере одним из названных видов значений, но лишь некоторые лингвистические формы обладают всеми четырьмя видами значений одновременно. Существуют лингвистические формы, обладающие только структурным значением (ср. флексию именительного падежа в русском языке) или структурным и сигнификативным (ср. словообразовательные аффиксы, морфемы типа «Пегас») и т. д. Это, однако, не создает никаких трудностей, поскольку разные виды значений являются объектами разных наук. Синтаксические и дифференциальные значения являются объектом структурной лингвистики. Сигнификативные и денотативные значения образуют предмет семантики [4]. Наконец, прагматические значения изучаются в прагматике и психолингвистике.

Относя тот или иной вид значения к предметной области определенной науки, мы еще ничего не говорим о возможных методах его изучения; методы некоторой науки могут быть использованы не только внутри нее самой, но и за ее пределами. В частности, как показал С. К. Шаумян [59], методы структурной лингвистики могут применяться в семантике.

В современной лингвистике к изучению значений были применены: 1) методы структурной лингвистики (коммутационный и компонентный анализ, метод субституции, метод НС и трансформационный метод); 2) психофизиологические методы; 3) методы математической логики; 4) методы логической семантики, 5) статистические методы. Некоторые методы (метод семантических множителей, метод тезаурусов), основанные одновременно на лингвистических, логических и математических идеях, выпадают из этой классификации.

О разных методах изучения значений можно говорить еще в одном смысле — с точки зрения задач, которые при этом ставятся. Если ставится задача найти эмпирический способ установления и количественного измерения значения, можно говорить об объективных (экспериментальных) методах; если ставится задача исследовать с помощью готового формального аппарата свойства некоторого множества значений при допущении, что элементы этого множества заданы извне, можно говорить о формальных методах [5].

Столь сложная ситуация приводит к тому, что при изложении существующих теорий значения и методов семантического исследования возникает определенная техническая трудность: за основу рубрикации текста можно взять или классификацию видов значений или любую из двух классификаций методов. Выбор решения произволен. В данном тексте основная рубрикация покоится на классификации значений, а более мелкая рубрикация — на классификации методов. Расположение материала определяется соображениями удобства и связности изложения, но отнюдь не важностью или ценностью того или иного метода.

Б. ИССЛЕДОВАНИЕ СТРУКТУРНОГО ЗНАЧЕНИЯ

Методы описания структурного значения так называемых грамматических морфем излагаются в многочисленных специальных работах (см., например, [98]). Поэтому в данном разделе целесообразно остановиться лишь на вопросе о том, каким образом и в какой мере методы,

[4] Некоторые логики (У. Куайн, Р. Карнап, Е. Бар-Хиллел) полагают, что должно быть две отдельные дисциплины: теория сигнификации («теория значения», по терминологии У. Куайна и Е. Бар-Хиллела) и теория денотации («теория референции», по терминологии У. Куайна и Е. Бар-Хиллела) [29], [65], [151].

[5] Отметим, что постановка обеих этих задач характерна для современного этапа развития лингвистики.

применявшиеся при изучении грамматики, могут быть использованы для описания структурных значений лексических морфем.

Многие исследователи (В. Н. Топоров, Р. Уэллс, З. Харрис, Ч. Хоккет и др.) считают, что подобное описание лексики или принципиально невозможно или, во всяком случае, практически чрезвычайно затруднено. В пользу этого тезиса выдвигаются следующие соображения: 1) между дистрибуцией лексического элемента и его значением нет взаимно-однозначного соответствия [99]; 2) лексический элемент нельзя представить как точку пересечения структурных отношений [53]; 3) классы лексических элементов, в отличие от классов грамматических элементов, бесконечны [103]; 4) лексические элементы, за редкими исключениями (ср. термины родства), не образуют парадигм [179].

Эти тезисы отнюдь не являются результатом исследования, а представляют собой гипотезу, которая нуждается в проверке. По существу данной гипотезы можно заметить следующее. 1) Неясно, что понимается под взаимно-однозначным соответствием между дистрибуцией лексического элемента и его значением; разумным образом этот тезис можно понять только в том смысле, что различия в дистрибуции лексических элементов не соответствуют различиям в их значении. Уточненный таким образом тезис и следует проверять. 2) По всей видимости, не всякий лексический элемент может быть представлен как «точка пересечения структурных отношений», но класс эквивалентных лексических элементов может быть представлен таким образом. Это значит, что возможно структурное описание лексики во всяком случае с точностью до класса[6]. 3) Нетрудно было бы доказать, что число лексических классов, определяемых на основе структурных признаков, конечно и не слишком велико; при этом то обстоятельство, что сами классы могут оказаться бесконечными, никак не может служить препятствием для структурного описания лексики. Из того, что класс бесконечен, следует только, что он не может быть задан по объему (перечислением), а должен быть задан по содержанию (т. е. совокупностью дифференциальных признаков). 4) Последний тезис в неявной форме предполагает, что только наличие парадигм позволяет представить множество языковых элементов в виде системы. Система есть там, где есть парадигмы, а структурное описание возможно лишь тогда, когда налицо некоторая система. Можно, однако, исходить из иного, более общего и точного понимания термина «система», предложенного И. А. Мельчуком, и считать, что некоторое множество объектов системно организовано, если число правил, с помощью которых могут быть построены эти объекты, меньше, чем число самих объектов. Тогда парадигматичность уже не может рассматриваться как единственный признак системной организации языковых фактов, и возникает задача поиска иных путей представления лексики как системы.

В свете сказанного мы бы считали метод описания структурных значений лексических элементов удовлетворительным, 1) если бы он обеспечивал построение классов эквивалентных лексических элементов, а также определение отношений и измерение расстояний между классами и 2) если бы число описывающих единиц (дифференциальных признаков

---

[6] В связи с этим возражением интересно отметить еще два факта: а) существуют морфемы, традиционно классифицируемые как грамматические, которые не допускают однозначных структурных определений (ср. морфемы времени в таком языке, как шведский); этот факт не считается достаточным основанием для того, чтобы отрицать возможность структурной грамматики; б) существуют лексические морфемы, которые вполне однозначно определяются через набор трансформаций и, следовательно, могут быть представлены как точка пересечения структурных отношений; такова английская морфема *take* в следующих фразах: *It* ***took*** *him a minute to wash—he* ***took*** *a minute washing—he* ***took*** *a minute to wash —washing* ***took*** *him a minute he* ***took*** *a minute in washing* (некоторые из этих трансформаций являются, должно быть, избыточными).

классов) было бы меньше числа описываемых единиц (т. е. самих лексических классов).

Эти или сходные идеи послужили отправной точкой при разработке проблем структурного описания лексики [3—5], [47], [52], [111], [121]. Для структурного описания лексики используется комплекс более или менее канонических методов структурной лингвистики, уточняемых в соответствии с обстоятельствами: компонентный анализ, метод НС и разработанный в последние годы трансформационный метод [77], [78], [100], [183].

Предполагается, что семантические различия и тождества до некоторой степени отражаются в синтаксических различиях и тождествах. Видимо, не всякое семантическое различие проявляется в каком-то синтаксическом различии, но всякому синтаксическому различию соответствует существенное семантическое различие. Если эта гипотеза верна, то классы лексических элементов можно описать через синтаксис — на основе дифференциальных синтаксических признаков, т. е. валентностей (синтаксических значений) лексических элементов. В один класс объединяются те лексические элементы, которые характеризуются одинаковым набором валентностей; при этом валентность лексического элемента определяется на основе дистрибутивных и трансформационных критериев. Поясним сказанное на примере. Класс лексических элементов, способных замещать позицию прилагательного в конструкции типа *воздух полон копоти*, *он достоин награды* (прилагательное подчиняет существительное в родительном падеже), отличается по своему синтаксическому значению (валентности) от класса лексических элементов, замещающих позицию прилагательного в конструкции типа *костюм мал мне*, *книга непонятна читателям*, *A эквивалентно B* (прилагательное подчиняет существительное в дательном падеже). Внутри второго класса можно выделить ряд подклассов в зависимости от того, как реагируют фразы с теми или иными лексемами на некоторые трансформационные преобразования. Оказывается, в частности, что особый подкласс образуют лексемы *подобный*, *тождественный*, *эквивалентный*, *синонимичный*, *равный* и т. д. в отличие от подкласса *мал*, *велик*, *понятен*, *известен*, *приличен* и т. д. Лексемы первого подкласса допускают преобразование типа *A тождественно (эквивалентно, синонимично, подобно, равно) B → A и B тождественны (эквивалентны, синонимичны, подобны, равны)*, а также преобразование типа *A эквивалентно (тождественно, синонимично, подобно, равно) B → B эквивалентно (тождественно, синонимично, подобно, равно) A*; ср. в противоположность этому лексемы второго подкласса (*костюм мне мал, ботинки велики ему, книга понятна читателям, поступок приличен офицеру*), не допускающие ни первого, ни второго преобразования. С другой стороны, лексемы второго подкласса допускают преобразование *костюм мал мне → костюм мал для меня, ботинки велики ему → ботинки велики для него, книга понятна читателям → книга понятна для читателей, поступок приличен офицеру → поступок приличен для офицера*, которое оказывается невозможным для лексем первого подкласса. Отметим еще, что если существительное в именительном падеже имеет форму единственного числа, то прилагательные класса {*эквивалентный*, *тождественный*, *синонимичный*, *подобный*, *равный*} обязательно требуют после себя существительного в дательном падеже, в то время как прилагательные класса {*мал*, *велик*, *понятен*, *приличен*} могут употребляться и в более простой конструкции — без существительного в дательном падеже, ср. *костюм мал*, *ботинок велик*, *книга понятна*, *поступок неприличен* при невозможности сказать **A тождественно (эквивалентно, синонимично, подобно, равно)* [7]. На основании

[7] Из этих примеров следует, между прочим, что категория числа, которая в известных нам системах синтаксической нотации никогда не отмечается как иррелевант-

этого различия в валентностях лексемы первого подкласса могут быть интерпретированы содержательно как выражающие двухместные предикаты, которые обладают свойством симметричности (если aRc, то cRa).

Каждый класс определяется структурно через набор дифференциальных синтаксических признаков (валентностей).

Внутри одного далее неразложимого класса лексические элементы попарно противопоставляются на парадигматической оси, т. е. находятся в отношении субституции или чередования (в смысле американских дескриптивистов). Иными словами, они обладают разными дифференциальными значениями (значимостями).

Поскольку каждый дифференциальный синтаксический признак имеет четкое структурное выражение, возникает основа для сопоставления лексических классов, т. е. определения отношений и измерения расстояний между классами [4], [5].

Можно надеяться, что описание лексики в терминах валентностей будет удовлетворять и второму сформулированному выше требованию. Во всяком случае, уже проведенные эксперименты показывают, что число дифференциальных синтаксических признаков действительно меньше, чем число лексических классов, которые можно построить на их основе.

В. ИССЛЕДОВАНИЕ ДЕНОТАТИВНОГО ЗНАЧЕНИЯ

Некоторые весьма скромные, но тем не менее заслуживающие упоминания результаты были получены при изучении денотативного значения психофизиологическими методами. Была разработана экспериментальная методика установления отношения слова к реальной вещи, основанная на механизме условного рефлекса. Интересные результаты были получены школой А. Г. Иванова-Смоленского, изучавшей взаимодействие первой и второй сигнальных систем [14], [28], [30], [38], [42], [60].

Основой всех этих работ является выдвинутое А. Г. Ивановым-Смоленским учение о процессах элективной иррадиации [27]. Сущность проведенных им опытов состоит в следующем. У испытуемого вырабатывается условный рефлекс (какая-либо двигательная реакция) на некоторый условный раздражитель, например вспышку зеленого света. Затем испытуемому предъявляются два словесных стимула (раздражителя), имеющих нечто общее с вещным стимулом: фраза *зеленый свет*, произносимая вслух экспериментатором, и та же фраза, написанная на табличке. А. Г. Иванову-Смоленскому удалось показать идентичность двигательных условных реакций на вспышку зеленого света, произнесение слов *зеленый свет* и предъявление их написанными на табличке. «Следовательно, данная реакция оказалась связанной не только с условным раздражителем, но и с его словесными символами — устным и письменным, не распространяясь, однако, на все другие слова (т. е. нося характер избирательной, с психологической точки зрения, — «смысловой» связи)» [27].

Все три условные связи образуют в мозговой коре единую динамическую структуру. Выработав условную реакцию на какой-либо один из раздражителей, связанных с такой структурой, экспериментатор получает сразу, без всякой предварительной подготовки, «как бы путем вне-

---

ная для синтаксической структуры предложения, является в действительности релевантной во всех тех случаях, когда мы имеем дело с двухместными симметричными предикатами. Фраза вида «существительное + связка + прилагательное», в которой прилагательное выражает двухместный симметричный предикат, может быть правильной лишь в том случае, если существительное имеет форму множественного числа, ср. *они тождественны (эквивалентны, подобны, синонимичны, равны)*. Множественное число скрывает в данном случае второй конец отношения.

запного замыкания» [27], ту же реакцию и на два других раздражителя. Это явление было названо «элективной иррадиацией».

В работах А. Г. Иванова-Смоленского и М. Марушевской [38] было установлено, что механизм элективной иррадиации действует не только при возникновении условного рефлекса, но и при его торможении. Выработав дифференцировку между зеленым и голубым световыми сигналами, они получали положительную реакцию на слова *зеленый свет* и тормозную — на слова *голубой свет*. «Иррадиация торможения, вызванная угашением ориентировочного рефлекса на непосредственный раздражитель, избирательно распространяется на его словесное обозначение, приводя к торможению ориентировочного рефлекса на этот словесный раздражитель» [38].

В описанных опытах механизм элективной иррадиации был использован для установления связей между раздражителями первой и второй сигнальных систем [8]. Иначе говоря, была выработана экспериментальная методика установления денотативного значения слов. В принципе с помощью описанного аппарата можно решать два основных вопроса, связанных с нахождением денотативных значений: устанавливать их тождества и различия. Положительная реакция на раздражитель является показателем тождества денотативного значения, а тормозная реакция — показателем различия денотативных значений.

Хотя принципиально эти вопросы, видимо, вполне разрешимы, перед исследователем денотативного значения стоят многочисленные практические трудности. Изученные до сих пор области (цвето- и звукообозначения вместе с реальными цветами и звуками) сравнительно легко исследуются с помощью изложенной методики. Неясно, однако, будет ли этот инструмент столь же эффективен в тех областях, в которых предъявление испытуемому реального референта сопряжено со значительными техническими трудностями (ср., например, референты глаголов).

Очевидно, что и сам измерительный инструмент, весьма грубый в его нынешнем виде, должен быть значительно усовершенствован, чтобы он мог точно и тонко индексировать различия денотативных значений. Усовершенствование измерительного инструмента будет, видимо, сочетаться с выработкой операционных (количественных или иных) индексов денотативного значения.

На этом мы заканчиваем рассмотрение методов исследования денотативного значения, хотя отнюдь не исчерпали темы. Дело в том, что в ряде случаев применение определенного аппарата к изучению сигнификативного значения ведет к установлению и денотативного значения. Было бы, однако, неразумно искусственным образом дробить единую систему понятий в угоду принятой схеме изложения материала. Поэтому в таких случаях весь относящийся к делу материал излагается в разделе о сигнификативном значении.

Г. ИССЛЕДОВАНИЕ СИГНИФИКАТИВНОГО ЗНАЧЕНИЯ

I. Методы логической семантики

1. Система Е. Бар-Хиллела и Р. Карнапа. В логической литературе семантика делится на дескриптивную и чистую [73]. Дескриптивная семантика — это исследование сигнификативных и денотативных значений естественных языков. Чистая, или логическая

[8] Он был использован и для установления связей между раздражителями второй сигнальной системы, т. е. между словами. Например, было обнаружено, что если у испытуемых выработана условная реакция на слово *хорошо*, она распространяется и на слова *отлично* и *плохо*. Однако этот факт имеет отношение к установлению семантических полей, т. е. сигнификативных, а не денотативных значений.

семантика занимается построением искусственных систем семантических правил, устанавливающих условия истинности языка-объекта. Примером такой системы является логическая семантика Р. Карнапа, состоящая из трех частей: 1) теории логической дедукции, 2) теории обозначения, 3) теории истины [73].

Логическая семантика имеет дело только с полностью формализованными языками [33], [160]. Поэтому ее понятия в настоящее время непригодны для анализа (сигнификативных) значений в естественных языках [25], [33], [76], [160] [9]. Тем более интересны некоторые работы Е. Бар-Хиллела и Р. Карнапа [29], [64], в которых делается попытка показать применимость основных понятий логической семантики к анализу значений в естественных языках [10].

Е. Бар-Хиллел исходит из известного положения логической семантики о том, что знакам языка могут быть приписаны два рода содержания: интенсионал (значение) и экстенсионал (референция) [11]. В соответствии с этим семантика мыслится Е. Бар-Хиллелом как дисциплина, включающая две теории с двумя системами понятий. «Одна теория имеет дело с интенсиональными или коннотативными аспектами языка и других знаковых систем, т. е. з н а ч е н и е м (в узком смысле слова); другая — с экстенсиональными или референционными аспектами» [64].

Теория значения рассматривает понятия логической истины, логической эквивалентности и синонимичности. Теория референции рассматривает понятия истины, равенства истинностных значений и коэкстенсивности.

Разработка первой теории целиком входит в задачи лингвиста. Лингвисту достаточно знать лишь интенсионалы выражений, чтобы установить, 1) является ли данное выражение логически истинным, 2) являются ли два произвольных выражения логически эквивалентными, 3) являются ли два произвольных выражения синонимичными.

Здесь весьма точно пересказаны концепции Р. Карнапа и У. Куайна [148]—[151] и выдвинуто утверждение, что с их помощью можно построить строгую систему семантических понятий для естественного языка. Фактически, как показал Н. Хомский, подвергнувший работу Е. Бар-Хиллела сокрушительной критике [76], Е. Бар-Хиллел вовсе не использует результатов Р. Карнапа или У. Куайна для анализа значений естественного языка. Дело в том, что Е. Бар-Хиллел предполагает наперед заданным весьма важное понятие интенсионала, как это делается в ранних работах Р. Карнапа. В логической семантике с ее ограниченным алфавитом символов эта процедура законна и легко формализуема, так как интенсионалы или классы синонимичных терминов могут быть заданы конечным и не слишком большим списком. Ее нельзя механически переносить на значения естественного языка. Действительно, чтобы показать верность следствия *это — красного цвета → это — не голубого цвета → это — окрашено*, необходимо фиксировать все эти высказывания в качестве постулатов или семантических правил данной системы. Следовательно, мы получаем характеристику синонимии в естественном языке, составляя полный список его синонимов (в число которых входят и синонимичные высказывания). Но тогда, как справедливо замечает Н. Хом-

[9] Ср. характерное высказывание Н. Хомского, может быть, даже излишне категоричное: «. . . более внимательное исследование допущений и объекта логического синтаксиса и семантики покажет, что надежда на применение результатов, достигнутых в этих областях, к решению лингвистических проблем иллюзорна» [76].

[10] К числу этих работ следует отнести и [65]. Выдвинутая здесь теория «семантической информации» представляет собой по существу язык узкого исчисления предикатов, более близкий к языкам логической семантики, чем к языку математической теории информации. Нам представляется, что весьма ограниченные средства теории «семантической информации» недостаточны для описания семантики естественных языков (ср., однако, замечания Е. В. Падучевой в [46]).

[11] Разъяснение этих понятий см. на стр. 106.

ский, вся дальнейшая теоретическая работа становится в высшей степени бессмысленной, так как такие списки, если бы они были возможны, решали бы все существенные вопросы [12].

Видимо, одна из основных задач лингвиста состоит в том, чтобы найти эмпирическую процедуру для установления интенсионалов выражений. В этом отношении Р. Карнап в своих последних работах [29] пошел дальше Е. Бар-Хиллела, поскольку он предложил именно такую процедуру. Она исключительно проста, если не примитивна, и состоит в том, чтобы путем опроса информанта, которого Р. Карнап условно называет «Карлом», выяснить интенсионалы употребляемых им выражений. Следует усомниться в том, что перед информантом можно ставить столь сложные задачи. С другой стороны, неясно, является ли эта процедура практически реализуемой, если учитывать, что в число «выражений» целесообразно включать не только морфемы, но и высказывания (во всех тех случаях, когда морфема многозначна). Даже если эта процедура практически реализуема, индивидуальное и независимое описание каждого интенсионала будет делом, практически весьма трудным и методологически неинтересным, поскольку оно должно быть чрезвычайно громоздким.

В структурной лингвистике была разработана процедура, которая дает возможность преодолеть эти трудности. Мы имеем в виду компонентный анализ, принципы которого были развиты Л. Ельмслевом [19], [102], Э. Сэпиром [159], З. Харрисом [97], [98] и другими исследователями. Компонентный анализ позволяет представить сигнификативное значение каждой морфемы как определенную совокупность нескольких элементарных независимых семантических единиц («фигур» в смысле Л. Ельмслева, «компонентов» в смысле З. Харриса), из конечного и относительно небольшого их набора. Иными словами, громадное число сложных единиц может быть сведено к ограниченному инвентарю элементарных семантических единиц, в терминах которого можно описать и все наличные в языке морфемы, и большое число несуществующих, но потенциально возможных морфем. Компонентный анализ играет исключительную роль, особенно при описании лексики, и ни одна адекватная теория смысла не может обойтись без процедур, обеспечивающих в той или иной степени компонентный анализ значений.

Применительно к изучению сигнификативного значения принципы компонентного анализа получили интересную разработку в ряде школ современной лингвистики (см. части II—V данного раздела).

2. И д е и А. Т а р с к о г о. Более плодотворный путь использования понятий логической семантики для анализа значений естественных языков указал А. Тарский [167]. Выдвинутый им принцип состоит в следующем: чтобы построить семантические понятия для естественного языка, необходимо сначала построить формализованный (информационно-логический) язык, служащий приближением к данному естественному языку. Можно построить серию таких моделей, в которой каждая по-

[12] Н. Хомский следующим образом суммирует существо рекомендаций Е. Бар-Хиллела: «Если лингвисту нужно установить или использовать каким-либо образом тот факт, что *окулист* и *глазник* являются синонимами, но его мучает совесть, он может избежать опасности ментализма, используя формальную процедуру провозглашения данного факта семантическим постулатом. А если он озабочен вопросом о том, почему *окулист — это глазник* является семантическим постулатом, а *Вашингтон — столица США* им не является, он может не сомневаться, что имеется строгая процедура для установления этого различия, а именно — включить первое высказывание в список семантических постулатов и не включать второго. Совершенно ясно, что такой случайный подход к проблеме классификации и характеристики элементов в тех или иных языках не поможет лингвистам, которых интересуют общие основания для установления этих элементов и отношений в каждом конкретном случае. Если у нас есть операционное описание синонимии. . ., формальная и строгая процедура составления произвольных списков не нужна; если же у нас нет операционного описания, эта процедура очевидно бесполезна, несмотря на свою буквальную формальность» [76].

семантика занимается построением искусственных систем семантических правил, устанавливающих условия истинности языка-объекта. Примером такой системы является логическая семантика Р. Карнапа, состоящая из трех частей: 1) теории логической дедукции, 2) теории обозначения, 3) теории истины [73].

Логическая семантика имеет дело только с полностью формализованными языками [33], [160]. Поэтому ее понятия в настоящее время непригодны для анализа (сигнификативных) значений в естественных языках [25], [33], [76], [160] [9]. Тем более интересны некоторые работы Е. Бар-Хиллела и Р. Карнапа [29], [64], в которых делается попытка показать применимость основных понятий логической семантики к анализу значений в естественных языках [10].

Е. Бар-Хиллел исходит из известного положения логической семантики о том, что знакам языка могут быть приписаны два рода содержания: интенсионал (значение) и экстенсионал (референция) [11]. В соответствии с этим семантика мыслится Е. Бар-Хиллелом как дисциплина, включающая две теории с двумя системами понятий. «Одна теория имеет дело с интенсиональными или коннотативными аспектами языка и других знаковых систем, т. е. з н а ч е н и е м (в узком смысле слова); другая — с экстенсиональными или референционными аспектами» [64].

Теория значения рассматривает понятия логической истины, логической эквивалентности и синонимичности. Теория референции рассматривает понятия истины, равенства истинностных значений и коэкстенсивности.

Разработка первой теории целиком входит в задачи лингвиста. Лингвисту достаточно знать лишь интенсионалы выражений, чтобы установить, 1) является ли данное выражение логически истинным, 2) являются ли два произвольных выражения логически эквивалентными, 3) являются ли два произвольных выражения синонимичными.

Здесь весьма точно пересказаны концепции Р. Карнапа и У. Куайна [148]—[151] и выдвинуто утверждение, что с их помощью можно построить строгую систему семантических понятий для естественного языка. Фактически, как показал Н. Хомский, подвергнувший работу Е. Бар-Хиллела сокрушительной критике [76], Е. Бар-Хиллел вовсе не использует результатов Р. Карнапа или У. Куайна для анализа значений естественного языка. Дело в том, что Е. Бар-Хиллел предполагает наперед заданным весьма важное понятие интенсионала, как это делается в ранних работах Р. Карнапа. В логической семантике с ее ограниченным алфавитом символов эта процедура законна и легко формализуема, так как интенсионалы или классы синонимичных терминов могут быть заданы конечным и не слишком большим списком. Ее нельзя механически переносить на значения естественного языка. Действительно, чтобы показать верность следствия *это — красного цвета → это — не голубого цвета → это — окрашено*, необходимо фиксировать все эти высказывания в качестве постулатов или семантических правил данной системы. Следовательно, мы получаем характеристику синонимии в естественном языке, составляя полный список его синонимов (в число которых входят и синонимичные высказывания). Но тогда, как справедливо замечает Н. Хом-

[9] Ср. характерное высказывание Н. Хомского, может быть, даже излишне категоричное: «. . . более внимательное исследование допущений и объекта логического синтаксиса и семантики покажет, что надежда на применение результатов, достигнутых в этих областях, к решению лингвистических проблем иллюзорна» [76].

[10] К числу этих работ следует отнести и [65]. Выдвинутая здесь теория «семантической информации» представляет собой по существу язык узкого исчисления предикатов, более близкий к языкам логической семантики, чем к языку математической теории информации. Нам представляется, что весьма ограниченные средства теории «семантической информации» недостаточны для описания семантики естественных языков (ср., однако, замечания Е. В. Падучевой в [46]).

[11] Разъяснение этих понятий см. на стр. 106.

ский, вся дальнейшая теоретическая работа становится в высшей степени бессмысленной, так как такие списки, если бы они были возможны, решали бы все существенные вопросы[12].

Видимо, одна из основных задач лингвиста состоит в том, чтобы найти эмпирическую процедуру для установления интенсионалов выражений. В этом отношении Р. Карнап в своих последних работах [29] пошел дальше Е. Бар-Хиллела, поскольку он предложил именно такую процедуру. Она исключительно проста, если не примитивна, и состоит в том, чтобы путем опроса информанта, которого Р. Карнап условно называет «Карлом», выяснить интенсионалы употребляемых им выражений. Следует усомниться в том, что перед информантом можно ставить столь сложные задачи. С другой стороны, неясно, является ли эта процедура практически реализуемой, если учитывать, что в число «выражений» целесообразно включать не только морфемы, но и высказывания (во всех тех случаях, когда морфема многозначна). Даже если эта процедура практически реализуема, индивидуальное и независимое описание каждого интенсионала будет делом, практически весьма трудным и методологически неинтересным, поскольку оно должно быть чрезвычайно громоздким.

В структурной лингвистике была разработана процедура, которая дает возможность преодолеть эти трудности. Мы имеем в виду компонентный анализ, принципы которого были развиты Л. Ельмслевом [19], [102], Э. Сэпиром [159], З. Харрисом [97], [98] и другими исследователями. Компонентный анализ позволяет представить сигнификативное значение каждой морфемы как определенную совокупность нескольких элементарных независимых семантических единиц («фигур» в смысле Л. Ельмслева, «компонентов» в смысле З. Харриса), из конечного и относительно небольшого их набора. Иными словами, громадное число сложных единиц может быть сведено к ограниченному инвентарю элементарных семантических единиц, в терминах которого можно описать и все наличные в языке морфемы, и большое число несуществующих, но потенциально возможных морфем. Компонентный анализ играет исключительную роль, особенно при описании лексики, и ни одна адекватная теория смысла не может обойтись без процедур, обеспечивающих в той или иной степени компонентный анализ значений.

Применительно к изучению сигнификативного значения принципы компонентного анализа получили интересную разработку в ряде школ современной лингвистики (см. части II—V данного раздела).

2. И д е и А. Т а р с к о г о. Более плодотворный путь использования понятий логической семантики для анализа значений естественных языков указал А. Тарский [167]. Выдвинутый им принцип состоит в следующем: чтобы построить семантические понятия для естественного языка, необходимо сначала построить формализованный (информационно-логический) язык, служащий приближением к данному естественному языку. Можно построить серию таких моделей, в которой каждая по-

[12] Н. Хомский следующим образом суммирует существо рекомендаций Е. Бар-Хиллела: «Если лингвисту нужно установить или использовать каким-либо образом тот факт, что *окулист* и *глазник* являются синонимами, но его мучает совесть, он может избежать опасности ментализма, используя формальную процедуру провозглашения данного факта семантическим постулатом. А если он озабочен вопросом о том, почему *окулист — это глазник* является семантическим постулатом, а *Вашингтон — столица США* им не является, он может не сомневаться, что имеется строгая процедура для установления этого различия, а именно — включить первое высказывание в список семантических постулатов и не включать второго. Совершенно ясно, что такой случайный подход к проблеме классификации и характеристики элементов в тех или иных языках не поможет лингвистам, которых интересуют общие основания для установления этих элементов и отношений в каждом конкретном случае. Если у нас есть операционное описание синонимии. . ., формальная и строгая процедура составления произвольных списков не нужна; если же у нас нет операционного описания, эта процедура очевидно бесполезна, несмотря на свою буквальную формальность» [76].

следующая модель сильнее предыдущей. Тогда описанием семантики естественного языка может служить перевод текстов этого языка на формализованные языки, из числа которых наиболее сильные могут быть описаны с помощью семантических правил логической семантики.

В недавнее время эта программа была частично реализована при описании семантики ряда узких предметных областей некоторых точных наук (в основном геометрии и химии). Семантическая организация текста (например, по химии), написанного на естественном языке, описывается на основе его соотношения с текстом на информационном языке (в данном случае на языке химических формул). Текст на естественном языке рассматривается в качестве переводимого, а текст на информационно-логическом языке — в качестве переводящего. Задача описания семантического строения текста решается следующим образом: 1) на основе однозначно соотносящихся с денотатами единиц переводящего текста, рассматриваемого в качестве эталона, в переводимом тексте выделяются некоторые элементарные (наименьшие) семантические комплексы; 2) описываются операции перевода (подробности см. в [32]).

II. Методы структурной лингвистики

1. Метод «коммутации» Л. Ельмслева. В работе, относящейся к 1957 г. [102], Л. Ельмслев попытался применить принципы глоссематики к изучению значений. Он наметил некоторые понятия «структурной семантики», которая строится им на основе соссюровского разграничения значения и значимости. Понятие значения относится к субстанции содержания языка, а понятие значимости — к структуре содержания.

Значимости — чисто дифференциальные, оппозитивные и негативные единицы, не несущие в себе ничего семантического. Они выявляются методом коммутационной проверки [13]. Этот метод позволяет установить [14], в каких случаях мы имеем дело с функцией коммутации (следовательно, с инвариантами или значимостями), а в каких — с функцией субституции (следовательно, с вариантами значимости). Так, в русском, английском и французском языках *он* (*he*, *il*) и *она* (*she*, *elle*) связаны функцией коммутации, а в финском, венгерском и китайском языках, где одно слово обозначает и *его* и *ее*, имеет место функция субституции.

Носителями значимостей являются знаки языка. На уровне некоторых знаков, например слов, состав единиц часто бесконечен; так, с точки зрения Л. Ельмслева, существительные нормально составляют открытый класс. Структурное описание может быть осуществлено только при условии, что найден некоторый способ сведения открытых классов к закрытым. Методом, с помощью которого осуществляется это требование, является компонентный анализ, проводимый на основе принципа коммутационной проверки [15]. Компонентный анализ исчерпывает задачи описания значимостей.

Инварианты, или значимости манифестируются в материальной форме, т. е. значениях [16]. Значения изучаются на двух уровнях — физическом и оценочном. Под физическим уровнем Л. Ельмслев понимает, по всей

---

[13] Та же методика используется в работах Е. Куриловича [31] и С. К. Шаумяна [55].

[14] При допущении, что денотаты рассматриваемых элементов известны исследователю.

[15] Последнее, конечно, необязательно. З. Харрис разработал оригинальную методику компонентного анализа морфем, основанную на чисто дистрибутивных процедурах.

[16] В работах Е. Найда [133] также исследуются значимость, значение и их взаимоотношения, хотя у него эти понятия обозначены другими терминами.

видимости, денотативное и сигнификативное значение; под оценочным — эмоционально-аффективную компоненту прагматического значения. Оценочный уровень представляется Л. Ельмслеву настолько важным, что в установленной автором иерархии он помещается над физическим уровнем. Изучение значения на этом уровне ближайшим образом связано с изучением общественных установлений данного лингвистического коллектива, т. е. его «социальной антропологией».

Таким образом, «структурная семантика» Л. Ельмслева включает в себя две теории: теорию значимости и теорию значения. По существу, однако, теория значимости не может быть частью семантики, потому что значимость, как подчеркивает Л. Ельмслев, не несет в себе ничего семантического.

В итоге от «структурной семантики» остается одна теория значения [17], которая носит у Л. Ельмслева незавершенный характер. Как подчеркивает сам Л. Ельмслев, значение в противоположность значимости относится к субстанции содержания. Со значениями мы имеем дело в речи, т. е. на эмпирическом уровне, и они должны изучаться эмпирически. Однако никаких процедур, удовлетворяющих этому требованию, Л. Ельмслев не предлагает.

Попытка разработать некоторую операционную методику описания значений была предпринята в исследовании А. Б. Долгопольского [18], наметившего путь построения «метаязыка для описания субстанции плана содержания». А. Б. Долгопольский исходит из принципиальных соображений Л. Ельмслева о функциях коммутации и субституции. В качестве объективного критерия оценки значения он вводит перевод с одного языка на другой. Сравнивая переводы одного и того же текста, он выделяет с помощью весьма тонкой методики («анализа по остатку») соотношения типа *рука — hand, arm*; *мыть, стирать — wash*. Различие *hand* и *arm*, релевантное в английском языке, нерелевантно в русском, а различие *мыть* и *стирать*, релевантное в русском, нерелевантно в английском. Таким образом предполагается выделить все семантические различия, релевантные хотя бы в одном языке [18]. Эти различия считаются элементарными единицами (семами) субстанции содержания. Предполагается, что если таким методом будет проанализирована достаточно большая группа языков, предпочтительно разного строя, можно будет построить исчерпывающее исчисление сем, достаточное для описания значений любого языка. Тогда каждое значение может быть представлено как некоторый набор сем [19].

Обсудим некоторые трудности, возникающие в исследовании такого рода.

1. При сравнении переводов текста на языки, принадлежащие резко отличающимся в культурном отношении обществам, трудно сегментировать высказывание и устанавливать соответствия, более дробные, чем соответствия целых высказываний [20]. Не всегда ясно, каким образом зна-

---

[17] Эта теория не заслуживает, видимо, наименования «структурной семантики».

[18] Задача описания всех семантических различий, релевантных хотя бы в одном языке, сводится к построению метаязыка, понимаемого как теоретико-множественная сумма этих различий. По всей видимости, такой метаязык был бы гораздо более сильным, чем это необходимо для мыслимых теоретических или практических целей. Разумнее, видимо, строить метаязык, достаточно сильный лишь для того, чтобы описать различия, релевантные во всех сравниваемых языках. Целесообразно при этом поставить задачу отождествления и разграничения не значений отдельных слов, но значений целых высказываний. Такой метаязык может быть понят как пересечение (в теоретико-множественном смысле) различий, обнаруживаемых в сопоставляемых языках (ср. «метод тезаурусов», раздел V).

[19] Ср. ниже метод «семантических множителей».

[20] Ср. мнение Э. Леннеберга, сравнивающего чукотский язык с английским: «Культурные и физические контексты чукотских высказываний, за немногими исключениями, несопоставимы с контекстами, в которых используется английский язык.

чение всего высказывания распределено между его элементами и какие аспекты действительности при этом выделяются (см. [152]).

2. В большинстве случаев оказывается, что данное высказывание может быть переведено многими способами (ср. [36]), и тогда, даже если нам удастся сегментировать его, данная сегментация является не единственно возможной в данном языке, но одной из нескольких вероятных. Необходимость введения вероятностных оценок неизмеримо усложняет задачи исследователя [21].

3. Если учесть еще, что 28 языков, привлеченных А. Б. Долгопольским для анализа, не создают, видимо, необходимой основы для построения семантического метаязыка требуемого типа (их число должно быть значительно больше) и что всю работу следует проделать на материале нескольких тысяч единиц, а не десятка слов, изученных в эксперименте А. Б. Долгопольского, станет совершенно ясно, что намеченная им интересная программа исследований неосуществима.

Идеи Л. Ельмслева развивал еще Р. Уэллс [179], пытавшийся моделировать с помощью метода коммутации некоторые семантические категории, в частности многозначность, омонимию и синонимию. Синонимия определяется им как соответствие одной единицы плана содержания более чем одной единице плана выражения. Многозначность — это такое соответствие одной единицы плана выражения более чем одной единице плана содержания, которое не является уникальным в данном языке. Так, английский глагол *take* ('брать', 'понимать') был бы признан многозначным, поскольку в английском языке имеется еще несколько глаголов с такой же или сходной комбинацией значений: *catch* 'хватать', 'понимать', *grasp* 'хватать', 'понимать' и др.[22] Омонимия определяется как такое соответствие одной единицы плана выражения более чем одной единице плана содержания, которое является уникальным, т. е. встречается в языке только один раз.

Неясно назначение этих определений. Если они формулируются в качестве строгой основы для опознания синонимии, многозначности и омонимии, то следует сказать, что они не достигают цели, поскольку элементы глоссематической терминологии отнюдь не придают им точности; если же все три категории предполагаются известными, то эти определения бесполезны.

В целом попытка построить понятия структурной семантики на основе идей Л. Ельмслева представляется нам неудачной.

2. Дескриптивные методы в семантике. Перенос методов дескриптивной лингвистики в область изучения сигнификативных значений был осуществлен в этнолингвистике.

В разнообразных исследованиях, которые объединяются в одно целое под названием этнолингвистики [54—56], [62], [66], [80], [91], [103], [104—106], [112], [119], [120], [121], [139], [177], развито несколько теорий значения. Общим для них является представление о языке как о своеобразном промежуточном мире (Zwischenwelt, Mittelwelt) между мышлением человека и действительностью, который навязывает мышлению определенный, своеобразный способ членения действительности. В остальном, однако, различные этнолингвистические теории значения существенным образом отличаются друг от друга; мы сосредоточиваем внимание лишь на одной из них, возникшей в связи с изучением терминов род-

Чукотское оружие, пища, нормы различного рода, природа, фауна и флора по большей части незнакомы обществам, говорящим на английском языке. Таким образом, практически не существует общего метаязыка (frame of reference), не существует основы для сегментного сравнения единиц двух языков. Перевод в этом отношении может быть лишь весьма грубым приближением к тому, что намеревался выразить автор» [122].

[21] Эта трудность была отмечена в частной беседе самим А. Б. Долгопольским.

[22] Аналогичный «критерий многозначности» предложил несколько раньше А. И. Смирницкий [51].

ства, так как она формулируется более строго, чем все другие [92], [114], [125].

Занимаясь терминами родства в различных языках, А. Кребер [114] обнаружил, что пары слов в этой терминологической системе могут быть противопоставлены друг другу по одному (или нескольким) из следующих восьми независимых признаков: 1) принадлежность к определенному поколению, 2) характер родственной связи (связь по прямой линии или побочная), 3) возраст внутри одного поколения, 4) пол родственника, 5) пол «эго», 6) пол лица, которое является одновременно родственником «эго» и другого лица, 7) кровное родство, 8) жизнь или смерть родственника. А. Кребер полагал, что выделенные им признаки универсальны и делают возможным описание терминов родства любого языка, хотя обычно тот или иной конкретный язык использует лишь некоторые из этих признаков.

На основе системы А. Кребера, уточняемой сообразно обстоятельствам, был выполнен ряд интересных исследований, из которых здесь будут рассмотрены работы У. Гуденафа [92] и Ф. Лаунсбери [125]. В названных работах нашли применение следующие принципы дескриптивной лингвистики: 1) противопоставление «эмических» и «этических» единиц, т. е. инвариантов и вариантов; 2) дистрибутивный анализ, особенно принцип дополнительной дистрибуции; 3) компонентный анализ (ср. [124]); 4) принцип построения парадигм.

Исследование строится обычно следующим образом: в том или ином языке на основании интуитивных соображений отбираются все термины родства, и каждый термин кодируется некоторым набором элементарных символов. Кодируя, например, английское слово *cousin*, исследователь учитывает все денотаты, которые могут быть обозначены этим словом (значения всех слов рассматриваются только по отношению к «эго»). Если примем следующие обозначения: «отец» — *F*, «мать» — *M*, «брат» — *B*, «сестра» — *S*, «сын» — *s*, «дочь» — *d* и условимся, что в нашей записи между каждым предыдущим и каждым последующим символом устанавливается отношение притяжательности, то английское слово *cousin* ('кузен', 'кузина') будет закодировано так: *FBs*, *FSs*, *MBs*, *MSs*, *FBd*, *FSd*, *MBd*, *MSd* [23].

Все эти символы представляют собой некоторые аллосемы (денотативные варианты сем или, по другой терминологии, семемы), так как в данном случае они не являются дифференциальными. Важно подчеркнуть, что в терминах этой системы можно построить конечное исчерпывающее исчисление аллосем, т. е. семантических вариантов.

Состав сем (дифференциальных семантических признаков) того или иного термина можно определить лишь после фиксации всех релевантных признаков, т. е. признаков, необходимых и достаточных для описания данной семантической системы. Каждый такой признак интерпретируется как некоторое независимое семантическое измерение в пространственной модели или как переменная величина, которая может принимать одно из фиксированного количества значений (обычно для одной переменной задаются два значения, так как описываемая процедура основана на двузначной логике). Измерения получают символические обозначения. Каждый термин системы рассматривается как пучок семантических компонентов или набор символов, по одному с каждого измерения. Так, для

[23] Существует много других денотатов слова *cousin*, например *FFSs*, *FMMBsd* и т. д., так как английское слово *cousin* присваивается денотатам, которые по-русски называются «двоюродным дядей», «двоюродной теткой», «двоюродным племянником», «двоюродной племянницей», «троюродным дядей», «троюродной теткой» и т. п. Однако по ряду существенных соображений предпочтительно анализировать слово *cousin* как два омонима. Поэтому перечисленные в сноске денотаты здесь нами не учитываются, а слово *cousin* пишется в дальнейшем *cousin*[1] (первый омоним).

описания термина *cousin* [1] на семическом уровне (в изложенном выше понимании; ср. аллосемическую запись) достаточно выделить следующие признаки: G — «поколение» (G_1 — младшее, G_2 — одно и то же, G_3 — старшее) и L — «характер родственной связи» (L_0 — родство по прямой линии, L_1 — побочное родство первой степени, L_2 — побочное родство второй степени). Тогда английское слово *cousin* [1] записывается как G_2L_2. В принципе можно последовательно, просто и исчерпывающе разграничить все термины данного семантического поля [24].

Набор терминов с дополнительным распределением сем составляет замкнутую группу слов, которая называется лексической парадигмой [25]. G_2L_2 входит в следующую парадигму: G_1L_0, G_1L_1, G_1L_2, G_2L_0, G_2L_1, G_2L_2, G_3L_0, G_3L_1, G_3L_2. Лексическая парадигма отличается от грамматической тем, что ни один из входящих в нее членов не сегментируется на морфологические части. Она сходна с грамматической парадигмой в том отношении, что порождает единицы, потенциально возможные в данной системе, даже если они не имеют плана выражения. Так, в приведенной выше парадигме лишь четыре термина выражаются отдельными словами, а пять терминов не имеют плана выражения:

G_1L_0 *child*	G_1L_2	G_2L_1 *sibling*	G_3L_0 *parent*	G_3L_2
G_1L_1	G_2L_0	G_2L_2 *cousin*	C_3L_1	

Лексические парадигмы можно перестраивать, получая все новые и новые группировки слов. Так, если мы введем в нашу систему признак S — «пол» (S_1 — мужской, S_2 — женский), то получим следующую парадигму:

$G_1L_0S_1$ *son*	$G_2L_2S_1$	$G_1L_1S_2$ *niece*	$G_3L_0S_2$ *mother*
$G_1L_1S_1$ *nephew*	$G_3L_0S_1$ *father*	$G_1L_2S_2$	$G_3L_1S_2$ *aunt*
$G_1L_2S_1$	$G_3L_1S_1$ *uncle*	$G_2L_0S_2$	$G_3L_2S_2$
$G_2L_0S_1$	$G_3L_2S_1$	$G_2L_1S_2$ *sister*	
$G_2L_1S_1$ *brother*	$G_1L_0S_2$ *daughter*	$G_2L_2S_2$	

Разные совокупности признаков, действующие в качестве операторов, дают разные парадигмы (классы слов). Это открывает перед этнолингвистикой широкие, пока еще не реализованные возможности. Отметим две из них.

Во-первых, может быть создана принципиально новая теория классов слов, отличающаяся от традиционной тем, что классы как таковые не будут включаться в описание. В описании будут сформулированы лишь законы построения классов. Для построения классов достаточно знать дифференциальные признаки, по которым один член данной системы может отличаться от другого. Задавая те или иные наборы признаков, исследователь сможет по мере необходимости строить классы большего или меньшего объема (ср. приведенные выше парадигмы), преобразовывать их по определенным правилам и т. д. В зависимости от заданного набора признаков один и тот же термин может входить в разные классы, причем число и состав классов для данной совокупности терминов будут всякий раз меняться. Получается гибкая и последовательная классификация, удачно названная И. А. Мельчуком «скользящей» [40].

Во-вторых, имея такую теорию классов, этнолингвистика сможет сделать новый шаг вперед на пути создания адекватной теории значения — разработать принципы семантической типологии. Обратимся вновь к на-

[24] Весьма существенно то обстоятельство, что для записи терминов на аллосемическом и семическом уровнях имеются два разных алфавита символов.

[25] Термин «лексическая парадигма» используется в лингвистике довольно давно, но разработка глубокого учения о лексической парадигме — достижение этнолингвистики.

шему примеру. Мы рассмотрели английское слово *cousin* [1]. В русском языке ему соответствуют два сложных термина, которые мы условно будем считать простыми: «двоюродный брат» и «двоюродная сестра». Признак пола родственника является в данном случае дифференциальным; то, что было аллосемой в английском языке, выступает в качестве самостоятельной семы в русском [26]. Следовательно, на аллосемическом уровне понятие «двоюродный брат» будет выражаться формулами *FBs*, *FSs*, *MBs*, *MSs*, но не формулами *FBd*, *FSd*, *MBd*, *MSd*, которые выражают понятие «двоюродная сестра». На семическом уровне эти два понятия выражаются следующими наборами сем: $G_2L_2S_1$ (двоюродный брат) и $G_2L_2S_2$ (двоюродная сестра). Весьма интересные различия обнаруживаются при сравнении второго омонима *cousin* [2] с соответствующими ему русскими терминами «двоюродный дядя», «двоюродная тетка», «двоюродный племянник», «двоюродная племянница», «троюродный дядя», «троюродная тетка», «троюродный племянник», «троюродная племянница». В русском языке в отличие от английского дифференциальными являются следующие три признака: 1) пола; 2) степеней побочного родства бо́льших, чем вторая; 3) поколения.

Таким образом, один и тот же набор аллосем организуется в двух языках в две различные системы.

Как показывают уже полученные результаты, термины родства хорошо описываются с помощью технического аппарата, созданного этнолингвистикой на основе принципов дескриптивной лингвистики. Возникает, однако, вопрос о его применимости за пределами такой относительно простой семантической системы, как термины родства. Мы называем эту систему простой, потому что 1) все исследуемые единицы данной совокупности однозначны, 2) они образуют замкнутую семантическую группу [27], 3) их денотаты дискретны (ср. термины родства в противоположность цветообозначениям), 4) тот или иной признак либо присутствует в данном элементе, либо отсутствует в нем [28]; постепенное нарастание признака невозможно (ср. признак пола в противоположность признаку размера).

Факты показывают, что возможности описанного нами аппарата ограничены [29]. Дело в том, что многие формы реального мира непрерывны [30]. Поэтому и в семантике, как это неоднократно отмечалось в литературе [2], [81], [125], [179], господствующим типом является не дискретная, а непрерывная система, своеобразный семантический континуум, в котором мало закрытых областей (т. е. полностью замкнутых, изолированных классов). В таких случаях основанные на двузначной логике детерминистские модели описанного типа оказываются слишком сильными. Указанные трудности можно преодолеть либо с помощью более слабых детерминистских моделей (ср. метод семантических множи-

[26] Ср. аналогичные ситуации в области фонологии.

[27] «Замкнутость» понимается здесь не в смысле конечности терминов (конечными являются не только такие группы слов, но и все множество лексических морфем языка), а в смысле четкой отграниченности от других семантических групп (полей). «Замкнутость» — это отсутствие плавных переходов.

[28] Это создает необходимую основу для применения двузначной логики.

[29] Это признают и авторы этнолингвистических работ. В частности, характерно, что Ф. Лаунсбери заканчивает свое обстоятельное исследование терминов родства следующим образом: «В некоторых областях лексики семантическая структура может быть настолько сложной, что становится невозможно или нецелесообразно подходить к ней. . . с аристотелевской логикой классов и принципом дифференциальной прагматической проверки в качестве основных орудий. Может возникнуть необходимость покинуть аристотелевскую дихотомию „*A* против не-*A*“ и ясное различие между „существенными“ и „случайными“ признаками. Вместо этих четких дихотомических делений могут быть введены непрерывные шкалы; может измеряться степень существенности того или иного признака; могут быть разработаны тесты, отличные от простого теста тождества и различия» [125].

[30] Имеются в виду индивидуальные объекты, но не ситуации.

телей, метод тезаурусов), либо с помощью многозначной логики и вероятных моделей (ср. статистические методы и метод семантического дифференциала).

III. Методы математической логики

1. Анализ семантического поля. Работа Дж. Гринберга. Подобно У. Гуденафу и Ф. Лаунсбери Дж. Гринберг исследовал семантическое поле терминов родства, но применил для его анализа более совершенный аппарат алгебры высказываний [93].

Анализ терминов родства дается во многих руководствах по математической логике (см. работы Р. Карнапа, А. Тарского и др.). Однако в большинстве случаев он предпринимается в чисто иллюстративных целях. Дж. Гринберг является, очевидно, первым и пока единственным исследователем, построившим законченную и экономную модель этого семантического поля. Его модель является аксиоматической; Дж. Гринберг вводит небольшое число первичных (неопределяемых) понятий и предложений, на основе которых определены все другие понятия и предложения системы. Построенная совершенно формально, модель интерпретируется затем для определенной предметной области — терминов родства.

Для того, чтобы уяснить принципы построения модели, достаточно рассмотреть ее фрагмент.

Вводятся следующие первичные отношения и объекты: P — отношение родителя к ребенку; xPy значит, что x является родителем y. M — отношение брака, упорядоченное следующим образом: в области отношения помещается символ, обозначающий мужа, а в противообласти — символ, обозначающий жену; xMy значит, что x является мужем y; m — класс мужчин. Предполагается заданным аппарат алгебры высказываний, т. е. логические связки $\vee$ (дизъюнкция; $x \vee y$ читается «x или y»); $\cdot$ (конъюнкция; $x \cdot y$ читается «x и y»); $\rightarrow$ (импликация; $x \rightarrow y$ читается «если x, то y» или «x имплицирует y» или «y следует из x»; $^-$ (отрицание; $\bar{x}$ читается «не-x» или «неверно, что x»); $\equiv$ (эквивалентность; $x \equiv y$ читается «x, если и только если, y»); кванторы всеобщности и существования, из которых мы используем только квантор существования $\exists$ ($\exists x$ читается «существует x, такой, что...»); кроме того, используется символ исчисления классов $\in$ (принадлежность к классу; $x \in C$ читается «x принадлежит C»); def. читается «по определению».

Определение: $f \equiv \bar{m}$ (f — класс женщин).
Предложение: $xMy \rightarrow (x \in m \cdot y \in f)$.
Предложение: $(xPy \cdot zPy \cdot z \neq x) \rightarrow (xMz \vee zMx)$ и т. д.
Определение: $Father \equiv$ def. x, y $(xPy \cdot x \in m)$;
$Mother \equiv$ def. x, y $(xPy \cdot x \in f)$.
$Brother \equiv$ def. x, y $((\exists u)(\exists v)\ uPx \cdot vPx \cdot uPy \cdot vPy \cdot x \neq y \cdot u \neq v \cdot x \in m)$;
$Sister \equiv$ def. x, y $((\exists u)(\exists v)\ uPx \cdot vPx \cdot uPy \cdot vPy \cdot x \neq y \cdot u \neq v \cdot x \in f)$;
$Husband \equiv$ def. x, y (xMy);
$Wife \equiv$ def. x, y (yMx);
$Son \equiv$ def. x, y $(yPx \cdot x \in m)$;
$Daughter \equiv$ def. x, y $(yPx \cdot x \in f)$.

Пусть B — *brother*, а S — *sister* (уже определенные понятия). Тогда $cousin \equiv$ def. x, y $((\exists u)(\exists v)(uBv \vee uSv) \cdot uPx \cdot vPy)$ и т. д.

Разработанный Дж. Гринбергом аппарат допускает строгое определение разных видов брака (моногамии, полигамии, полиандрии), сложных систем родства самых разнообразных обществ (в частности, ряда индейских, африканских и полинезийских племен), построение сложных

высказываний о допустимых и запрещаемых брачных связях и т. д. В этом отношении он превосходит, видимо, аппарат этнолингвистики; он, однако, не имеет каких-либо преимуществ перед аппаратом этнолингвистики в том отношении, что возможности его применения за пределами простых семантических полей столь же ограничены (см. стр. 119).

2. Семантический анализ слов. Попытка Дж. Гринберга является единственной в своем роде. В большинстве случаев идеи и методы математической логики используются для анализа сигнификативного значения отдельных слов, а не семантических полей.

Работы, посвященные семантическому анализу отдельных слов, опираются на разработанное в современной математической логике учение о языке-объекте и метаязыке [33], [160].

Описываемый язык — это язык-объект; описывающий язык выступает по отношению к нему в качестве метаязыка. Принципиально мыслимо неограниченное число метаязыков. Метаязык, непосредственно описывающий первичный язык-объект, является метаязыком первой степени; последний в свою очередь выступает в качестве языка-объекта по отношению к метаязыку второй степени и т. д. Всякий метаязык должен быть сильнее описываемого языка объекта.

Семантический метаязык создается путем анализа значений некоторых слов естественного языка. Б. Рассел показал [158], что в естественном языке скрыто по крайней мере два языка, которые принадлежат разным уровням [31]: язык-объект, включающий, главным образом, предметные слова, имена собственные, имена предикатов и т. п., и метаязык, состоящий из словесных аналогов логических связок, кванторов и других логических констант. Из слов языка-объекта можно составлять так называемые атомарные предложения, в которых отражаются некоторые эмпирические данные; они, следовательно, могут быть доказаны или опровергнуты ссылкой на эти данные. Из атомарных предложений строятся с помощью кванторов, логических связок и т. п. молекулярные предложения, теория которых является наиболее полной и разработанной частью логики. Все самые трудные проблемы связаны, по словам Б. Рассела, с атомарными предложениями (аналог ядерных предложений в трансформационной грамматике).

Каждое атомарное предложение содержит в идеале конечное число имен собственных (имен предметов) и одно и только одно слово, являющееся именем предиката (Б. Рассел приводит в качестве примера одноместный предикат «желтый», двухместный предикат «раньше», трехместный предикат «между»). Имена собственные отличаются от имен предикатов тем, что встречаются в атомарных предложениях любой формы, в то время как имя предиката встречается только в таких атомарных предложениях, которые имеют соответствующее число имен собственных (так, если предикат двухместный, то атомарное предложение должно содержать два имени собственных).

Б. Рассел показал далее, что принадлежащие метаязыковому уровню слова типа «все», «некоторые», «многие», «никакие», «не», «и», «или», «истинный», «ложный», неопределенный и определенный артикли встречаются только в молекулярных предложениях и не встречаются в атомарных. Он дал поучительный семантический анализ этих слов [32], ко-

---

[31] Именно этим объясняется возможность большого числа разнообразных семантических парадоксов, вроде парадокса лжеца, в обычной речи: слова «истинный» и «ложный», участвующие в парадоксе лжеца, принадлежат не языку-объекту, а метаязыку. Ср. глубокие идеи А. Тарского по этому вопросу [167], [168].

[32] Аналогичный анализ логических констант дается в большинстве руководств по математической логике и в некоторых специальных исследованиях (см., например, [75]).

торый привел его к нетривиальным выводам о синтаксическом и семантическом строении предложений некоторых типов. Так, предложения типа *I saw a man* ('я видел человека') кажутся на первый взгляд атомарными (в трансформационной грамматике Н. Хомского, например, они трактуются как ядерные). Поскольку, однако, они не могут быть эмпирически опровергнуты (как не могут быть эмпирически доказаны предложения, содержащие слово «все»), их следует трактовать как молекулярные предложения вида *I saw x, and x is human* (' я видел *x*, и *y* — человек'). Таким образом, вопрос об их эмпирическом подтверждении или опровержении снимается; для молекулярных предложений существен лишь вопрос о том, построено ли данное предложение в соответствии с правилами вывода.

Изложенные здесь положения (особенно, деление слов на имена собственные, имена предикатов, имена кванторов и других логических констант, основанное на понятии языка-объекта и метаязыка и связанное с теорией атомарных и молекулярных предложений) представляются нам исключительно важными, несмотря на их общий характер; они дают лингвисту ключ к решению многих синтаксических и семантических вопросов.

IV. Метод семантических множителей

Судя по немногочисленным публикациям лаборатории машинного перевода при Первом МГПИИЯ [21], [22], [23], [36], предложенный ею метод семантических множителей не вышел еще из стадии первоначальной разработки. Однако уже сейчас можно сказать, что его перспективное значение весьма велико, так как с его помощью можно будет, по всей видимости, решать гораздо более широкий круг задач, чем это первоначально предусматривалось.

Авторы исследования о семантических множителях (А. К. Жолковский, Вяч. В. Иванов, Н. Н. Леонтьева, Ю. С. Мартемьянов, В. Ю. Розенцвейг, Ю. К. Щеглов) выдвинули в качестве отправной точки развиваемой ими системы положение о том, что слово не является элементарной смысловой единицей языка. Это в особенности касается тех слов (глаголов, прилагательных, предлогов), которые организуют фразу (в логических языках им соответствуют имена отношений или предикатов) [33]. Сложное значение каждого такого слова разлагается на так называемые семантические множители, т. е. элементарные смыслы, являющиеся дифференциальными семантическими признаками слова (ср. фонемы и их дифференциальные признаки). На первом этапе работы дифференциальные семантические признаки выделялись интуитивно на основании сравнения синонимичных высказываний. Так, из сравнения высказываний *он недомогает — он нехорошо себя чувствует* извлекаются элементарные семантические множители: «отрицание» (ср. *не-* в слове *нехорошо*), «положительность» (ср. *-хорошо*) и «ощущение» (ср. *чувствует*). Тогда значение слова *недомогать* записывается как комбинация трех множителей.

Некоторые члены группы (см. [21]) применили в последнее время еще одну методику выделения семантических множителей: на основе интуитивного анализа толкового словаря (был использован, в частности, «Словарь русского языка» С. И. Ожегова) был составлен относительно небольшой список слов (до 2000), в терминах которого можно описать

[33] «К семантическим множителям имеет смысл обращаться в тех случаях и в той мере, в какой необходимо изобразить двухсторонние отношения, образующие своего рода костяк семантической структуры фразы, и прочие слова, смысл которых регулярно соотносится со смыслом других слов. . . В той мере, в какой слова не представляют интереса ни с одной из этих точек зрения, их запись в словаре может быть сведена к порядковому номеру. В особенности это относится к словам конкретной лексики, т. е. таким, где в первую очередь важны не смысл с его структурой, а предметная отнесенность» [23].

многие другие слова языка. Предполагается, что существует способ сведения списка в 2000 слов к еще более ограниченному инвентарю семантических множителей, так как эти 2000 слов не являются семантически независимыми. Некоторые из них непосредственно анализируются на элементарные семантические множители; некоторые другие рассматриваются как комбинации уже определенных слов и элементарных семантических множителей; сочетание множителей «предположение»+«истинное» (допускать, что x истинно) определяет значение слова *догадываться*; добавление к нему множителя «способность» дает значение слова *интуиция* и т. д. Весьма важен следующий принцип: при анализе смысла слова на семантические множители учитывается все множество реальных ситуаций, представителем которых является данное слово. Например, при анализе слова *прокурор* должны быть исследованы все стороны судопроизводства, с которыми регулярно связываются функции прокурора. Тогда каждый семантический множитель можно рассматривать как элемент некоторой ситуации, описываемой средствами языка, а самое ситуацию — как упорядоченное множество таких элементов.

С помощью семантических множителей записывается смысл целых фраз, предварительно анализируемых синтаксически по особому алгоритму, разработанному Ю. С. Мартемьяновым.

Сопоставляемая фразе семантическая запись может быть преобразована с помощью некоторых фиксированных правил. При этом оказывается, что одна и та же «приведенная» семантическая запись может быть сопоставлена весьма различным фразам (в некоторых экспериментах число фраз, имеющих данную семантическую запись, доходило до 400). Таким образом устанавливается смысловая инвариантность фраз (с точностью до преобразований). Кроме того, на основе «приведенной» семантической записи выявляются закономерности сочетания семантических множителей в пределах фразы.

Уже первые работы лаборатории показали, что в терминах небольшого числа семантических множителей можно описать весьма большое число реальных слов языка. Это объясняется тем, что разнообразие слов связано не только с разницей в их смысле, но и с рядом других обстоятельств: 1) разные слова могут быть синтаксически обусловленными вариантами выражения одного и того же смысла (ср. временное значение деепричастия, союза *когда*, предлога *при*); 2) разные слова могут быть лексически обусловленными вариантами выражения одного и того же смысла (ср. *оказывать влияние*, но *производить впечатление*); 3) некоторые пары слов, выражающих двухместные отношения, представляют один и тот же смысл в разных направлениях (ср. *иметь* и *принадлежать*, *раньше* и *позже*).

Поскольку, таким образом, число семантических множителей намного меньше числа сложных словесных значений, представляется возможным построить семантическое исчисление, в котором в терминах небольшого числа неопределяемых элементарных семантических множителей и аксиоматически вводимых первичных предложений можно определить большое число сложных значений и вывести большое число сложных предложений.

Значение нового подхода отнюдь не исчерпывается тем, что он создает принципиальную возможность экономного и строгого описания сигнификативных значений естественного языка в виде исчисления. Система искусственных семантических множителей, когда она будет создана, вместе с необходимыми элементами синтаксиса может рассматриваться как семантический метаязык, гораздо более эффективный, чем естественный язык, поскольку он будет менее избыточен и в значительной мере формализован. Его сила заключается и в его общности: поскольку семантические множители могут считаться общими для достаточно большого числа языков, метаязык приобретает весьма важные черты универсализма.

Эти два качества, вместе взятые, позволяют рассматривать его как своего рода приближение к логически совершенному языку (в смысле Л. Витгенштейна и Б. Рассела). Глубокий смысл исследований в области семантических множителей состоит именно в том, что они в конечном счете могут привести к созданию богатого и строгого логического языка.

Не лишено интереса и то обстоятельство, что рассмотрение значения в терминах семантических множителей позволяет по-новому поставить традиционные проблемы многозначности, синонимии и контекста; в частности, контекст понимается как определенная комбинация множителей. Новый подход позволяет, далее, вскрыть семантическое тождество слов, принадлежащих к разным грамматическим (функциональным) классам, а также тождество лексических и грамматических морфем [34]. Он позволяет, наконец, обнаружить неизвестные традиционной семасиологии отношения, для которых у нее нет даже названий (ср. ряд *иметь — дать — получить — взять — потерять — сохранить*).

По-видимому, вместе с методом тезаурусов, разработка которого продвинулась несколько дальше, метод семантических множителей представляет собой одно из самых перспективных и плодотворных направлений современной семантики.

V. Метод тезаурусов

Метод тезаурусов разрабатывается Кембриджской лингвистической группой в составе М. Мастерман, С. Джоунз, М. Кея и некоторых других исследователей [83], [109], [110], [127], [128], [132], [182] [35]. Свою основную задачу они формулируют следующим образом: «В принципе нас не интересует, каковы общие признаки находящегося перед нами текста и других текстов на том же языке. Мы требуем прежде всего и главным образом обнаружения того, что о т л и ч а е т данный текст; каков признак, делающий его д а н н ы м текстом, а не каким-либо другим; каково, наконец, заключающееся в нем сообщение» [110].

1. Т е о р и я. Теоретические положения Кембриджской группы выросли из их практической работы в области машинного перевода, в основу которой был положен идеологический словарь П. Роже. Как известно, словарь П. Роже, названный тезаурусом, представляет собой собрание слов английского языка, разбитых на ряд тематических классов; классы в свою очередь разбиваются на подклассы, каждый из которых представляет собой объединение нескольких более мелких тематических групп (по терминологии кембриджцев — «тем»). Допускается пересечение классов. Всего в словаре имеется 6 больших классов, 24 подкласса и 1000 тем. Примеры тем: «целое», «часть», «начало», «конец», «непрерывность», «дискретность», «время», «переход», «изменение», «завершение», «успех», «неудача».

Темы понимаются как типовые семантические контексты для слов, а класс семантических контекстов слова — как множество тем, которые могут быть ему приписаны [132]. Элементарный однозначный контекст (тема) принимается в качестве первичного неопределяемого понятия семантической теории, развиваемой Кембриджской лингвистической группой [128].

---

[34] Традиционные представления, связанные в особенности с именем А. И. Смирницкого, долгое время препятствовали осознанию этого очевидного факта, хотя в логической литературе он был отмечен давно. Ср. следующее высказывание Б. Рассела: «Части речи, как они выступают в грамматике, не имеют особого внутреннего отношения к логическому синтаксису. „До“ предлог, а „предшествовать“ — глагол, но они значат одно и то же» [158].

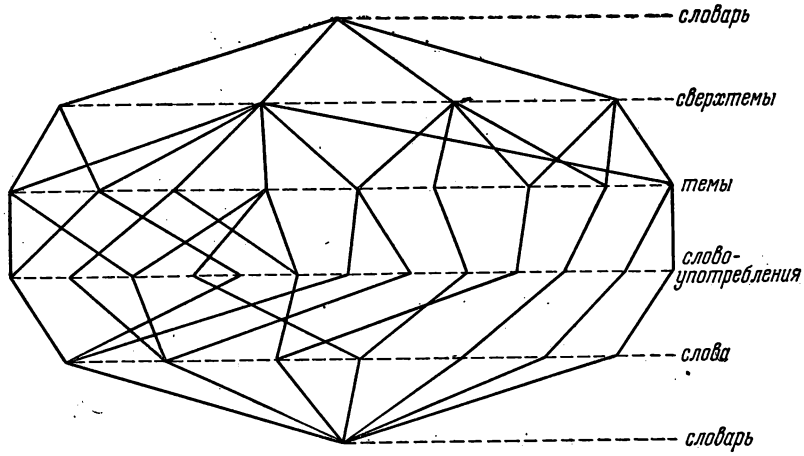
[35] Содержательный обзор работ Кембриджской группы был дан А. К. Жолковским [20].

Каждому словоупотреблению приписываются индексы тем, к которым оно относится. Поэтому каждое слово можно охарактеризовать с помощью определенного набора индексов, причем наборы индексов для разных слов, как правило, различны. Например, слово *finish* ('заканчивать') имеет следующий набор индексов тем: 44) разделение, 67) конец, 70) дискретность, 106) время, 113) мгновенность, 140) изменение, 142) прекращение, 292) прибытие, 729) завершение.

Количество словоупотреблений в принципе бесконечно. Однако число первичных внелингвистических ситуаций (тем, контекстов), которые люди должны уметь различать для успешного общения, конечно. Может показаться, что всякий раз, когда слово получает новое употребление, мы имеем в виду новую внелингвистическую ситуацию. Это впечатление неверно. Фактически мы используем бесконечно разнообразные синонимические средства выражения для обозначения небольшого набора уже имеющихся и известных нам внелингвистических ситуаций [127] (ср. стр. 123).

Итак, тезаурус классифицирует, во-первых, семантические контексты, во-вторых — слова. Семантическая классификация контекстов — это разбиение всего словаря на сверхтемы (классы тем), темы и т. п. Семантическая классификация слов — это приписывание данному слову индексов всех тех контекстов (тем), в которых оно типично встречается в разных случаях употребления. Эти две классификации могут быть представлены как булева структура. Естественно, поэтому, что математическим аппаратом «структурной семантики» в смысле кембриджцев стала теория структур.

Все сказанное выше получает при геометрической интерпретации следующий вид (дается принципиальная схема, отнюдь не воспроизводящая богатства фактически отражаемых в тезаурусе отношений):



Итак, каждое слово индексируется в соответствии с тем, в какие группы, темы и сверхтемы оно входит, причем общее число индексов может быть довольно большим (для предлога *in* 'в' оно достигает 200).

Важно подчеркнуть, что каждое слово сопоставляется со всем словарем и получает определение на основе полного объема значений, выражаемых в данном языке. Поэтому оказывается возможным разграничить очень большое число слов. Практически не разграничиваются лишь слова, отличающиеся только денотатами, например специальные термины какой-либо узкой предметной области. Для различения таких слов используются чисто условные числовые индексы [132] (ср. сноску 33).

Метод тезаурусов позволяет записать с помощью индексов тем не только значение слова, но и значение предложения. При записи предложения учитываются следующие два обстоятельства: 1) одинаковым образом записываются грамматическая и семантическая информации; это связано с тем, что одно и то же содержание выражается один раз (или в одном языке) лексически, а другой раз (или в другом языке) грамматически, ср. следующие синонимичные предложения: *я прочитал книгу — я закончил чтение книги*, которые должны быть записаны одинаковым образом [110] (ср. стр. 123); 2) смысловая информация, составляющая содержание данного предложения, складывается не из всех индексов тем, которые приписаны словам, но лишь из некоторых. Считается, что все темы, которые существенны для данного сообщения, встречаются в нем не менее двух раз. Поэтому все индексы, которые ни разу не повторяются в предложении, отбрасываются.

В составленной таким образом записи сохраняется не вся смысловая информация, которую при желании можно было бы извлечь из предложения. Однако сохраненная информация оказывается вполне достаточной для того, чтобы отличить данное предложение от другого, не совпадающего с ним по смыслу.

Законченный тезаурус рассматривается в качестве интерлингвы — логического языка-посредника, который достаточно слаб для того, чтобы быть универсальным, т. е. годиться для описания семантики любого языка, и достаточно силен для того, чтобы быть алгоритмически описанным или построенным.

Мы не будем подробно комментировать этих положений; все то, что было сказано на стр. 123 — 124 относительно метода семантических множителей, целиком относится и к методу тезаурусов.

2. Эксперимент: автоматический способ построения тезауруса. Этот вопрос рассматривается в работе С. Джоунз [109]. Предлагаемая автором процедура построения тезауруса разбивается на два этапа: 1) построение рядов, т. е. классов эквивалентных словоупотреблений, 2) построение понятийных групп, т. е. классов рядов (соответствующих «темам»). Первая операция не является автоматической и поэтому проводится самим исследователем. Вторая операция имеет вполне автоматический характер и поэтому проводится машиной.

Рядом называется конечный класс элементов, члены которого взаимозаменимы в позиции n предложения S_n. Предложение определяется как конечная последовательность элементов, ограниченная точечными знаками и имеющая свойство, называемое «плоем». Последнее понятие ясно не определено. Можно лишь предполагать, что плоем называется семантический инвариант предложений типа *это был большой ученый, это был выдающийся ученый, это был крупный ученый* (слова *большой, выдающийся, крупный* образуют ряд).

Практически, однако, процедура замещения не используется при построении рядов. Ряд строится на основе анализа того, как слово толкуется в словаре. Из словарных определений исследователь извлекает синонимичные группировки, рассматриваемые как ряды.

Для построения группы (темы) делается следующее допущение: разные словоупотребления одного слова (ср. *легкая ноша, легкая жизнь, легкое вино, легкий характер, легкое поведение*) семантически близки друг другу. На основе этого допущения составляются группы. Если у нас имеются ряды словоупотреблений 1) a, b, c, 2) e, b, c, 3) i, b, c [36], то они объединяются в одну группу на том основании, что все три ряда со-

[36] Ср., например: *Он ударяет (бьет, колотит) его; Сердце стучит (бьется, колотится) в груди; Его трясет (бьет, колотит) лихорадка.*

держат словесные знаки b и c. Факт совпадения одного знака может быть чисто случайным и поэтому недостаточным для установления групп. Чем больше число словесных знаков, общих для нескольких рядов, тем больше семантическое сходство этих рядов. Мера семантической близости двух рядов вычисляется по формуле:

$$S = \frac{\text{число знаков, общих для двух рядов}}{\text{число различных знаков в двух рядах}}.$$

Для установления групп используется несколько формул. Приведем одну из них. Некоторая совокупность C считается группой, если а) для всех $x \in C \cdot y \in C$, $S_{x,y} > \Theta$, где Θ — некоторый эмпирически установленный порог; б) не существует C', включающей C, такой, что C' удовлетворяет требованию а).

Процедура построения групп для ограниченного материала была представлена в виде алгоритма, запрограммирована и опробована на машине. Эксперимент дал хорошие результаты; выработанная машиной классификация групп вполне соответствовала интуитивному представлению о семантических сходствах и различиях классифицированных словоупотреблений.

VI. Статистические методы

Нам известны статистические исследования сигнификативного значения двух родов: 1) исследования, характеризующие значение как определенное семантическое место во фразе; 2) исследования, характеризующие значение как определенную точку семантического поля.

1. Значение как семантическое место во фразе. Эксперимент, типичный для исследований такого рода, описан в [61]. Испытуемым предъявляются короткие и длинные предложения с пропуском одного слова (обычно существительного), ср. *по . . . промчался грузовик* и *я долго сидел около открытого настежь . . . и сильно простудился*. Им предлагается указать первые восемь слов, которые, по их мнению, могут стоять во фразе на месте прочерка, причем эти восемь слов должны быть записаны в порядке убывающей вероятности их появления в данном месте. Готовый список возглавляют, таким образом, слова, которые считаются наиболее вероятными для данной позиции. В результате эксперимента было обнаружено, что 1) слова, называвшиеся в качестве наиболее вероятных для некоторой позиции, действительно являются наиболее вероятными для нее по данным статистических исследований; 2) наибольшая степень согласия в ответах испытуемых характерна для тех слов, которые воспринимаются как наиболее вероятные для данной позиции; 3) пропуск в более длинном предложении замещается меньшим числом разных слов (при учете отчетов всех испытуемых), чем пропуск в коротком предложении.

Ценность этих, в общем, тривиальных результатов состоит в том, что они экспериментально подтверждают обусловленность значения лингвистической формы (в частности, слова) местом, которое она занимает в цепочке синтагматически связанных с нею форм, и длиной этой цепочки. К сожалению, более интересный вопрос о мере этой обусловленности, предполагающей трудоемкое исследование условных вероятностей, в упомянутой нами работе не ставился.

2. Построение и анализ семантических полей. Принципиально новых классификационных результатов не принесло и исследование А. Я. Шайкевича, предложившего статистическую методику выделения семантических полей языка [58]. Тем не менее оно замечательно в двух отношениях: 1) семантические связи слов в системе (семантическом поле) устанавливаются на основе их синтагматических связей в тексте; 2) семантические связи слов устанавливаются на чисто

статистической основе, т. е. совершенно объективно. В этих условиях совпадение полученных А. Я. Шайкевичем результатов с интуитивно сложившимися традиционными классификациями представляет принципиальный интерес.

А. Я. Шайкевич предположил, что семантические связи слов должны проявляться, в частности, в их относительно частом совместном появлении в текстах. Поэтому, изучая статистические корреляции слов в тексте, можно делать выводы относительно их семантических связей. Для проверки этой гипотезы было проведено следующее исследование: в произведениях некоторых английских поэтов были подсчитаны прилагательные. Была вычислена вероятность совместного появления пары прилагательных в одной строке стихотворного произведения при предположении их случайного распределения в тексте. Затем выяснилось, сколько раз данное прилагательное фактически встретилось с каким-либо другим прилагательным в одной строке. Если полученная цифра была намного больше той, которая ожидалась (на основе предположения о случайном распределении), констатировалось наличие семантической связи между данной парой прилагательных [37]. В результате была составлена карта семантических полей языка, четко фиксирующая поля цветообозначений, размера и ряд других.

В заключение следует сказать, что описанные нами работы являются единственными в своем роде; широкое исследование семантического аспекта языка с помощью статистических методик еще не началось [38].

VII. Экспериментальные психофизиологические методы

Здесь будут рассмотрены две экспериментальные методики: 1) разработанная К. Ноблом методика измерения значения, основанная на механизме психической ассоциации; 2) предложенная О. С. Виноградовой и А. Р. Лурия экспериментальная психофизиологическая методика построения и анализа семантических полей, основанная на механизме условного рефлекса.

1. Измерение значения. Эксперимент К. Нобла. В статье «Анализ значения», опубликованной в 1952 г. [135], американский психолог К. Нобл предложил экспериментальную методику измерения и индексирования сигнификативного значения, которая получила широкое распространение и была использована для решения ряда психологических и психолингвистических задач [71], [136], [153], [162], [181].

Экспериментальный статистический индекс значения был получен К. Ноблом следующим образом. Из списка слов, характеризующихся различным порядком приближения к структуре английского языка, было выбрано 96 двухсложных слов. Среди них было 20% бессмысленных звукосочетаний, 35% малочастотных слов (меньше одного на 4 000 000) и 45%, частотных слов (больше одного на 1 000 000). Они предъявлялись по одному группе испытуемых в составе 119 человек. Испытуемые должны были в течение 60 *сек* записать все слова, которые возникают у них в сознании при предъявлении данного ключевого слова (стимула). При выводе количественного индекса значения не учитывались следующие виды письменных ответов испытуемых: 1) неразборчивые ответы, 2) повторяющиеся ответы (*ОТЕЦ — мать, мать, мать. . .*), 3) нерелевантные ответы, которые распадаются на свободные ассоциации (весьма туманное понятие) и звуковые ассоциации. Все остальные виды

[37] А. Я. Шайкевич провел еще одно исследование, совпадающее по своим принципиальным установкам с описанным опытом.

[38] Весьма интересные семантические алгоритмы Ю. А. Шрейдера, основанные в определенной мере на числовых методах, стали известны автору после завершения данной работы и поэтому не нашли в ней отражения.

ответов считались «допустимыми». Значение m_s данного словесного стимула вычисляется по формуле

$$m_s = \frac{1}{N} \sum_{i=1}^{N} R_s,$$

где s — словесный стимул (раздражитель), R_s — допустимый письменный ответ испытуемого, а N — число испытуемых. К. Нобл показал, что значение m возрастает по мере перехода от бессмысленных звукосочетаний к словам и от малочастотных слов к частотным. К. Нобл конструировал индекс m как показатель значения; однако в ходе работы он отождествил значение с осмысленностью или, точнее, с количеством осмысленных ассоциатов слова [39]. В результате оказалось, что бессмысленные звукосочетания тем не менее имеют значение, отличное от нулевого, потому что у них есть осмысленные ассоциаты (допустимые ответы). С другой стороны, в работе К. Нобла можно обнаружить целые группы слов, очевидным образом различающихся по значению, но одинаково частотных (т. е. в конечном счете одинаково осмысленных) и получающих поэтому один и тот же индекс m, ср. *jelly* ('желе'), *heaven* ('небеса'), *jewel* ('бриллиант').

Методика К. Нобла приводит, таким образом, к весьма приблизительным количественным оценкам сигнификативного значения. Это дало повод некоторым исследователям, в частности Ч. Осгуду и его сотрудникам, утверждать, что показатель m вообще не индексирует никаких аспектов значения [147]. Интересно, однако, что в новых исследованиях самого К. Нобла, Дж. Дженкинса и У. Рассела, Артура и Каролины Стаатс [162] было экспериментально доказано, что между индексом m и мерой (прагматического) значения D, которую предложили Ч. Осгуд и его сотрудники, имеется довольно высокая степень корреляции, т. е. что m и D связаны некоторой зависимостью, хотя и не идентичны друг другу. Это объясняется тем, что m индексирует какие-то аспекты сигнификативного значения, хотя, видимо, не все. Действительно, слово вместе со своими ассоциатами образует семантическое поле. Хотя значение слова не равно сумме значений его ассоциатов, оно включает в себя некоторый компонент, который является пересечением его значения со значениями ассоциатов, т. е. инвариантом данного семантического поля. Показатель m полезен в той мере, в какой он индексирует именно этот аспект сигнификативного значения.

Эта сторона дела может быть серьезным образом усовершенствована, если за основу подсчета брать не «допустимые ответы испытуемых», а другой тип ответов. Можно исходить из следующих эмпирических фактов. Экспериментально установлено, что предъявление испытуемым некоторого слова вызывает у них с х о д н ы е ассоциативные реакции, благодаря чему в их ответах на данное слово имеются с о в п а д а ю щ и е слова. Здесь уместно изложить результаты классической работы Г. Кента и А. Розанова [141]. Они составили список из 100 общеупотребительных слов английского языка (существительных и прилагательных) и получили набор ассоциатов для каждого слова этого списка от 1000 испытуемых, представлявших все слои населения США. Большинство слов-стимулов (т. е. 100 существительных и прилагательных) имело ограниченное число ассоциатов, повторявшихся от испытуемого к испытуемому. Так, на слово *needle* ('иголка') 712 испытуемых ответили, в частности, словами *thread* ('нитка'), *pin* ('булавка'), *sharp* ('острый'), *sew* ('шить'),

[39] Вполне сознательно это сделано в коллективной работе последователей К. Нобла [71].

sewing ('шитье'). В ответах 152 испытуемых содержалось слово *sharp*. С другой стороны, в ответах было много индивидуальных ассоциатов, т. е. таких, которые встретились в ответах только одного испытуемого, например, *weapon* ('оружие').

Мы допускаем, что ключевое (предъявляемое) слово входит в семантические поля слов, встретившихся в ответах испытуемых более одного раза, и не входит в семантические поля слов, встретившихся в ответах всего один раз (в этом случае их появление объясняется случайной или индивидуальной ассоциацией). На этом основании можно предложить экспериментальный индекс, фиксирующий наличие инварианта некоторого семантического поля у значения предъявляемого слова. Пусть N — число испытуемых, n_i — число ответов, содержащих некоторое слово w_i, w_0 — предъявляемое слово, а $M_{w_0}(w_i)$ — мера его вхождения в семантическое поле слова w_i. Тогда искомый инвариант для слова w_0 вычисляется по формуле

$$M_{w_0 w_{(i)}} = \frac{n_i - 1}{N},$$

В числителе стоит $n_i - 1$, так как в силу наших допущений n_i должно быть не менее 2, чтобы данное слово w_0 могло войти в семантическое поле слова w_i. При $n=1$, $n - 1=0$, и вопрос о вхождении слова w_0 в семантическое поле слова w_i тем самым снимается. В частном случае значения этой дроби будут равны $\frac{N-1}{N}$ и $\frac{1}{N}$. Таким образом, индекс $M_{w_0}(w_i)$ фиксирует не только наличие у данного слова w_0 семантического инварианта некоторого семантического поля, но показывает и степень вхождения данного слова в данное семантическое поле, т. е. то, в какой мере инвариант семантического поля характеризует значение w_0. Иными словами, фиксируется и наличие релевантного признака, и его степень.

Если вычислить значение дроби для всех ответов, встречающихся хотя бы два раза, можно установить все семантические поля, в которые входит данное слово. Слово *needle*, например, вошло бы, в частности, в семантические поля слов *pin sharp*, *sewing* и т. п.

Такая постановка вопроса привела бы к более точному анализу сигнификативного значения слова как такового и семантических полей, в которые оно в той или иной мере входит.

2. Построение семантических полей. Работа О. С. Виноградовой и А. Р. Лурия. Наиболее интересные результаты в этой области исследований были получены О. С. Виноградовой, А. Р. Лурия и Н. А. Эйслером [11], [12] [40].

Их методика является психофизиологической, так как она основана не на регистрации чисто психических реакций испытуемого на предъявление того или иного стимула, а на регистрации реакций его кровеносных сосудов. Была создана аппаратура, способная строго объективно фиксировать три типа реакций кровеносных сосудов испытуемого на предъявление определенного стимула: оборонительную, ориентировочную и нулевую (отсутствие реакции).

Первым этапом процедуры является выработка у испытуемого условного рефлекса на ключевое слово будущего семантического поля, например слово *скрипка*. Для этого ему предъявляются нейтральные слова *стена*, *карандаш*, *улица* и т. п., которые сначала вызывают ориентировочную реакцию. Через некоторое время она угасает. Тогда испытуемому много-

[40] Представляют интерес написанные под руководством А. Р. Лурия и О. С. Виноградовой дипломные работы студенток МГУ Г. Д. Пестовой «Объективное исследование систем смысловых связей» (1958) и Л. Н. Шуструйской «Взаимодействие обобщимых и необобщимых систем словесных связей» (1958). Автор благодарит проф. А. Р. Лурия за предоставленную ему возможность познакомиться с этими работами.

кратно предъявляется слово *скрипка*, причем предъявление слова каждый раз сопровождается подачей небольшой дозы электрического тока, который вызывает оборонительную реакцию кровеносных сосудов. Через некоторое время у испытуемого вырабатывается условный оборонительный рефлекс на слово *скрипка*, т. е. он начинает реагировать на данное слово, не подкрепленное электрическим током, точно так же, как он реагирует на ток.

На втором этапе, после того, как у испытуемого образуется условный оборонительный рефлекс на слово *скрипка*, ему предъявляются другие слова (без тока). Оказывается, что существует большая группа слов помимо слова *скрипка*, предъявление которых испытуемому сопровождается точно такой же реакцией. Таковы слова *смычок*, *струна*, *скрипач*. Существует, далее, группа слов, предъявление которых испытуемому вызывает лишь ориентировочную (более слабую) реакцию. Таковы слова *контрабас*, *виолончель*, *рояль*, *музыка*, *оркестр*, *соната*. Существует, наконец, большая группа слов, не вызывающих у испытуемого никакой реакции. Таковы слова *корова*, *стена*, *карандаш*, *улица*.

Характер реакции кровеносных сосудов головы и руки испытуемого на предъявление того или иного слова является, таким образом, показателем смысловой близости слов [41]. Поэтому на третьем этапе работы создаются условия для построения вокруг слова *скрипка* некоторого семантического поля. При этом слова первой группы (*смычок*, *скрипач*, *струна*) образуют центральный круг поля. Слова второй группы (*флейта*, *контрабас*, *рояль*, *музыка*, *оркестр*, *соната* и др.) образуют его периферию. Нейтральные слова (*корова*, *печка*, *стена*, *карандаш*, *улица*), предъявление которых не вызывает у испытуемого никакой реакции, располагаются за пределами поля.

Построив семантическое поле вокруг слова *скрипка*, экспериментатор может перейти к фиксированию других семантических полей. С этой целью предъявление испытуемому какого-либо нейтрального слова, например слова *корова*, сопровождается током. В результате этого оборонительная реакция на существительное *скрипка* угасает, и вырабатывается оборонительная реакция на новое ключевое слово. Она распространяется на все связанные с ним слова (*теленок*, *лошадь*, *стадо* и др.). Таким образом выделяется новое семантическое поле.

В ходе построения семантических полей интересно решаются некоторые вопросы, связанные с многозначностью. Предположим, что экспериментатор строит семантическое поле вокруг слова *скрипка*. В него должно быть включено и слово *труба* (музыкальный инструмент). Оно действительно включается в это поле, если вводится после слова *барабан*. Если же оно вводится вместе с существительными типа *камин*, *печка*, оно не вызывает никакой реакции и, следовательно, не попадает в семантическое поле слова *скрипка*. Два разных значения слова *труба* оказываются, таким образом, четко отграниченными друг от друга.

Конечно, описанная методика построения семантических полей еще далека от совершенства. Она дает самое общее и грубое представление о смысловом сходстве и различии слов, поскольку семантические связи классифицируются всего на три группы: сильные, слабые, нулевые [42]. Ее сильные стороны состоят в том, что семантическое поле выделяется вполне объективно на основании показаний аппарата, регистрирующего реакции кровеносных сосудов испытуемого на предъявление тех или иных слов.

[41] Было экспериментально показано, что слова, близкие к ключевому слову по звучанию, не вызывали никакой реакции. Это было подтверждено и в опытах Л. А. Шварца [60].

[42] Большее число семантических связей было обнаружено в интересной, но менее точной работе Дж. Флавелл и Е. Флавелл [87], где найдено, что степень близости двух слов зависит от характера логических отношений между понятиями, которые они выражают.

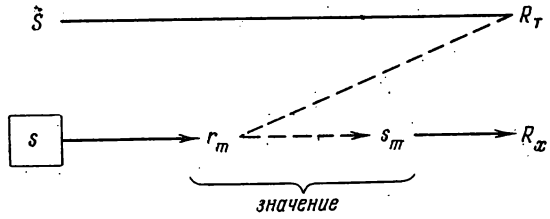
Д. ИССЛЕДОВАНИЕ ПРАГМАТИЧЕСКОГО ЗНАЧЕНИЯ

Наиболее важными и, видимо, единственными заслуживающими упоминания работами в области исследования прагматического значения являются работы Ч. Осгуда и его школы. В 1952 г. Ч. Осгуд опубликовал свою первую статью об измерении значения методом семантического дифференциала [140]. Она положила начало серии экспериментальных исследований, ход которых регулярно освещался в печати [116], [141—146], [166], [169]. Этот этап исследований завершился созданием обобщающего коллективного труда «Измерение значения» [147] и весьма важного «Атласа семантических профилей» [107], [108].

С момента первых публикаций к ним было привлечено внимание лингвистов и специалистов смежных областей [49], [74], [82], [90], [143], [175], [176]. Методы Ч. Осгуда и его школы получили широкое признание в американской психолингвистике, психологии и психометрии [69], [79], [117], [123], [126], [137], [156], [161—164], [180], а его учение стало одним из весьма влиятельных направлений современной экспериментальной семантики. Мы рассмотрим сначала теоретический и экспериментальный аппарат, разработанный Ч. Осгудом и его сотрудниками (раздел I). Затем будут описаны основные эксперименты и их результаты (раздел II). В заключение будут высказаны некоторые критические замечания по поводу работ Ч. Осгуда и его школы (раздел III).

I. Теоретический и экспериментальный аппарат

1. Теоретический анализ значения. В книге «Измерение значения» авторы приводят следующую схему, в которой подытожено их понимание значения:



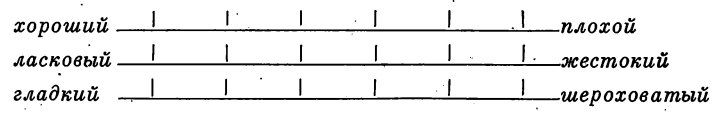
$\dot{S}$ обозначает денотат, который служит для организма безусловным стимулом (раздражителем) и в некоторой ситуации регулярно и надежно вызывает предсказываемую модель поведения R_T (безусловный рефлекс). Стимул $\boxed{s}$ — это лингвистический знак (условный раздражитель), отличный от денотата $\dot{S}$, но воздействующий на организм одновременно с ним. Поскольку $\boxed{s}$ не совпадает с $\dot{S}$, он вызывает не поведение R_T, но «репрезентативный опосредствованный процесс» r_m — — → s_m, который является лишь частью поведения R_T (ср. поведение, вызываемое у человека видом скорпиона и восклицанием «Скорпион!»). Между R_T и r_m (внутренней реакцией организма на знак) устанавливается условно-рефлекторная связь. Внутренняя реакция r_m вызывает «самостимуляцию» s_m (внутреннее раздражение организма), которая в свою очередь может вызывать некоторое поведение R_x. Отношение r_m к s_m («репрезентативный опосредствованный процесс») авторы считают значением знака.

Разные стимулы вызывают реакции, отличающиеся друг от друга качеством и интенсивностью (ср. «мороз», «холод», «прохлада» — «тепло», «жара», «зной»).

Авторы полагают, что «репрезентативный опосредствованный процесс» включает денотативное и прагматическое значения, которые и являются объектом измерения [43].

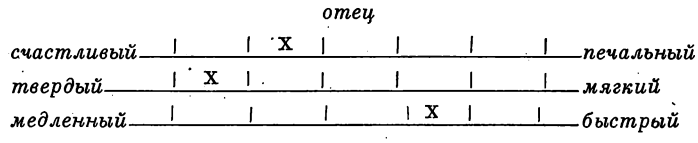
2. Операционный индекс значения. Ч. Осгуд допускает, что значение слова может быть определено как некоторая точка ряда непрерывных шкал, определяемых двумя полярными терминами. В соответствии с этим он строит свой экспериментальный инструмент для измерения значений, называемый семантическим дифференциалом (СД).

СД — это метод количественного и качественного индексирования значения с помощью двухполюсных шкал с семью делениями на каждой, задаваемых парой антонимичных прилагательных [44], например:



Если x — левый член шкалы, а y — правый, то деления слева направо содержательно определяются для испытуемых следующим образом: 1) очень x, 2) вполне x, 3) немного x, 4) ни x, ни y (и x, и y), 5) немного y, 6) вполне y, 7) очень y.

Испытуемому предъявляются те или другие понятия (*отец*, *симфония*, *леди*, *герой*, *грех*, *булыжник*, *нищий*), которые он в соответствии с этой инструкцией должен поместить на одно из делений каждой шкалы.



Каждое деление имеет номер (употребляются две системы нумерации: от +3 до —3 с 0 посредине и от 1 до 7 с 4 посредине.

В самых первых опытах Ч. Осгуда было обнаружено, что между различными шкалами имеется значительная степень соответствия [45]. Так, суждения на шкалах *хороший — плохой*, *справедливый — несправедливый*, *ласковый — жестокий*, *честный — бесчестный* коррелируют друг с другом, но не зависят от суждений на шкалах *длинный — короткий*, *мокрый — сухой* и т. п. На этой основе Ч. Осгуд постулировал свое семантическое пространство (далее именуемое СП). Группы эквивалентных (т. е. зависимых друг от друга) шкал объединяются в один фактор или одно семантическое измерение, независимое от другого семантического измерения, которое также задается группой эквивалентных шкал. Ограниченное число таких измерений, допускающих различение наибольшего числа понятий, может быть использовано для построения СП. Ч. Осгуд допустил, что все измерения пересекаются в одной точке (центре шкал, т. е. отметке 0 или 4) под прямым углом друг к другу.

Значение каждого слова может быть определено операционно и квантитативно как вектор [46], выходящий из центра СП (точки пересечения измерений).

---

[43] В разделе III мы покажем, что фактически измеряются лишь некоторые аспекты прагматического значения. Поэтому нельзя считать точными все те термины, которые содержат прилагательное «семантический» (СД, семантическое пространство).

[44] Одним из первых исследователей, использовавших для измерения значения шкалы такого типа, был К. Мозьер [131], опиравшийся на список антонимичных прилагательных К. Хаагена [96].

[45] Статистически такие соответствия оцениваются с помощью особой методики, которая носит название корреляционного анализа (см. об этом [89]).

[46] Вектором называется прямая линия, задаваемая исходной точкой, длиной и направлением.

Ч. Осгуд, Дж. Сучи и П. Танненбаум следующим образом связывают теоретическое определение значения с его операционным индексом: «Точка пространства, которая служит для нас операционным определением значения, имеет два существенных свойства — направление и расстояние от центра семантического пространства. Мы можем отождествить эти свойства соответственно с качеством и интенсивностью значения» [147].

В заключение данного параграфа подчеркнем, что введение семизначных шкал, с помощью которых можно индексировать не только качество, но и интенсивность значения (т. е. признак и его степень), позволяет осуществить важный переход от двузначной логики к многозначной логике.

II. Основные эксперименты и результаты

1. Построение СП. Для выяснения структуры СП, т. е. для установления числа независимых измерений, было проведено три основных эксперимента, статистические данные которых обрабатывались на электронной вычислительной машине [47].

В первом эксперименте 100 испытуемых давали суждения о 20 понятиях на 50 шкалах. Всего было получено $100 \times 20 \times 50 = 100\ 000$ суждений («куб данных»). Факториальный анализ [48] показал, что в матрице корреляций представлено четыре независимых фактора: 1) фактор оценки, имеющий большие веса на таких шкалах, как *хороший — плохой, красивый — безобразный, чистый — грязный, ласковый — жестокий* и др.; 2) фактор силы, имеющий большие веса на таких шкалах, как *большой — маленький, сильный — слабый, тяжелый — легкий, толстый — тонкий* и др.; этот фактор сильно переплетается с первым: оба имеют значительные веса на шкалах *громкий — тихий, твердый — мягкий, храбрый — трусливый* и др.; 3) фактор ориентированной активности, имеющий большие веса на таких шкалах, как *быстрый — медленный, активный — пассивный, жаркий — холодный* и др.; третий фактор также переплетается с первым: оба имеют значительные веса на шкалах *молодой — старый, бешеный — спокойный, напряженный — расслабленный*; 4) четвертый фактор, объясняющий всего около 2% различий, не удалось интерпретировать разумным образом.

Для второго эксперимента были использованы те же 50 шкал, но в качестве испытуемых было взято 40 новых студентов, и задача была изменена. Испытуемым предъявлялись пары шкал типа *острый — тупой, расслабленный — напряженный*, и они должны были обвести кружком то слово второй пары, которое представляется им более близким по значению к главному слову первой пары (*острый*). В примере, приведенном выше, в большинстве случаев обводится слово *напряженный*. Почти полное согласие достигается и в паре шкал *прямой — кривой, благородный — скотский*. В противоположность этому в паре шкал *свежий — засохший, длинный — короткий* одинаково часто обводится и *длинный*, и *короткий*. Иными словами, шкалы *прямой — кривой* и *благородный — скотский* зависимы друг от друга, т. е. представляют одно измерение. Шкала *свежий — засохший* независима от шкалы *длинный — короткий*; в данной паре представлены два разных измерения.

Описанный эксперимент вскрыл ту же факториальную структуру, т. е. те же семантические измерения, что и первый опыт.

В третьем эксперименте по сравнению с первым было увеличено до 76 число шкал. Он подтвердил наличие первых трех факторов. В дополнение к ним были выделены еще пять малосущественных факторов, которые мы

[47] Факт машинной обработки экспериментальных данных имеет принципиальное значение для лингвистики вообще.

[48] О факториальном анализе см. [89], а также [44], [45].

не будем здесь анализировать. Важно подчеркнуть, что все дальнейшие эксперименты (за исключением [107], [108]) основаны на трехмерной модели СП.

В идеале каждое измерение должно быть представлено одной шкалой, которая полностью совпадает с данным фактором и, следовательно, абсолютно независима от каких-либо других шкал. Таких шкал, однако, авторам практически не удалось обнаружить. Поэтому в большинстве экспериментов каждый фактор представлен тремя-четырьмя шкалами, максимально зависящими друг от друга и минимально зависящими от других шкал.

2. Определение значений слов (размещение понятий в СП). Значение слова определяется операционно как набор цифровых данных, характеризующих положение слова на том или ином факторе [49]. Ниже приводится таблица цифровых факториальных данных для 10 понятий в оценке одного испытуемого (табл. 1).

Таблица 1

	Понятия / Факторы	Оценка	Сила	Активность
A	*quicksand* 'зыбучий песок'	—3	3	—3
B	*white rose buds* 'бутоны белых роз'	3	—3	—3
C	*death* 'смерть'	—3	1	—3
D	*hero* 'герой'	3	3	3
E	*methodology* 'методология'	3	0	2
F	*fate* 'судьба'	—1	0	—2
G	*virility* 'мужественность'	2	3	2
H	*gentleness* 'мягкость'	2	—2	—3
I	*success* 'успех'	2	3	2
J	*sleep* 'сон'	2	0	—2

Позднее Дж. Дженкинс, У. Рассел и Дж. Сучи [102] вычислили средние значения для 360 понятий в оценке 540 человек на 20 шкалах, представлявших 6 факторов.

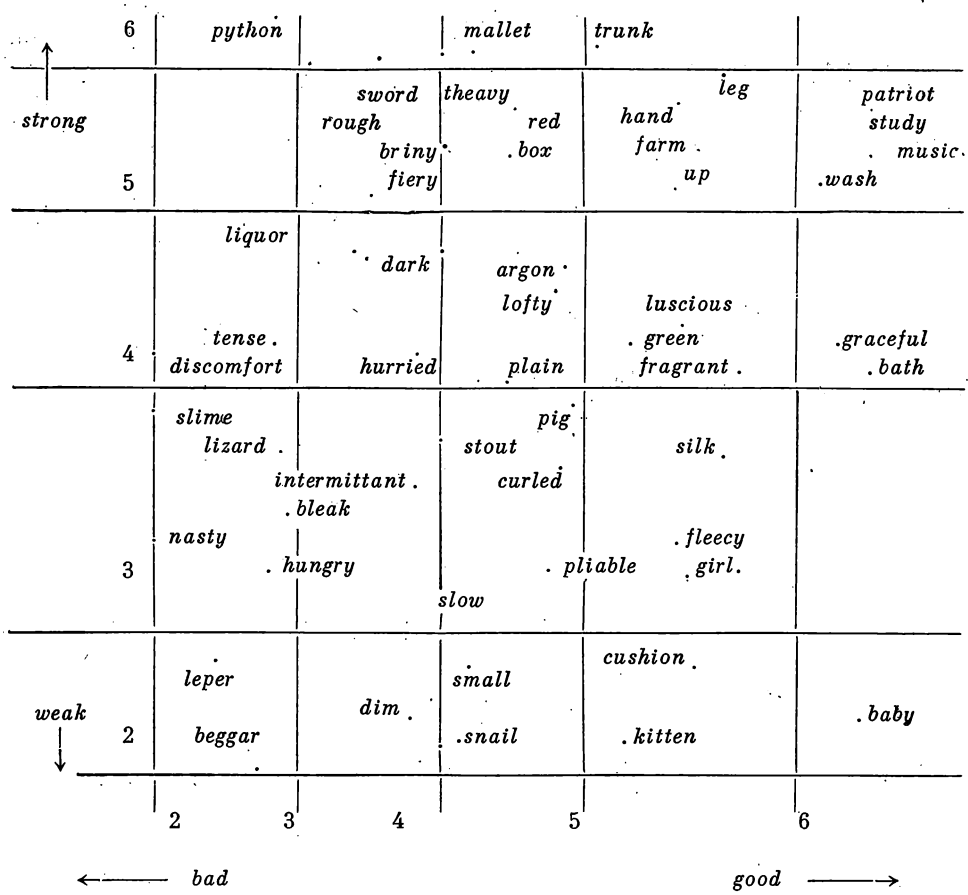
Цифровые данные могут быть интерпретированы геометрически. Одна из удобных геометрических интерпретаций состоит в том, что значение слова определяется как точка в СП, задаваемая ее координатами на нескольких факторах. В качестве примера приведем одну из таблиц «Атласа семантических профилей» [74] — табл. 2.

Возможна еще одна геометрическая интепретация цифровых данных — они изображаются в виде «профилей» на некоторых шкалах или факторах (ср. «семантический профиль» слова «отец» на стр. 133). Позднее Дж. Дженкинс, У. Рассел и Дж. Сучи составили такие профили для 360, оцененных 540 испытуемыми на 20 шкалах [107].

В большинстве экспериментов по измерению значений, описанных Ч. Осгудом и его сотрудниками, в качестве «понятий» брались существительные и реже прилагательные. Возникает естественный вопрос о том, сможет ли метод СД «взять» значения слов других грамматических классов (например, наречий, глаголов, предлогов). В литературе о СД и теории СП

[49] Каждое значение может быть представлено как пучок компонентов — по одному с каждого измерения. Эмпирическая гипотеза о компонентной структуре значений была экспериментально подтверждена в ряде остроумных опытов, поставленных Артуром и Каролиной Стаатс [161], [163], [164].

Таблица 2



высказывались вполне определенные сомнения по этому поводу. В частности, Дж. Кэрролл [74] усматривает один из существенных недостатков метода СД в том, что он не приспособлен для измерения значений предлогов, глаголов, наречий и т. п. Экспериментальные исследования, опубликованные после выхода в свет «Измерения значения» [107], [161] [164], [180], не подтвердили этих опасений. Методом СД были измерены значения всех этих классов слов, включая предлоги, причем полученные результаты оказались отнюдь не бессмысленными с точки зрения СД. Так, испытуемые последовательно и надежно оценивали глагол *eat* ('есть') как нечто доброе, своевременное, вкусное, округлое, сильное, обычное, красивое и т. д. С нашей точки зрения, это свидетельствует не о силе, а о слабости СД, который измеряет не сигнификативное и денотативное значения, но лишь некоторые (оценочные) аспекты прагматического значения.

3. Вычисления расстояний между значениями. Созданная Ч. Осгудом и его последователями модель дает возможность измерять расстояние между любой парой значений в СП. Оно вычисляется по формуле

$$D_{il} = \sqrt{\sum_j d_{il}^2}$$

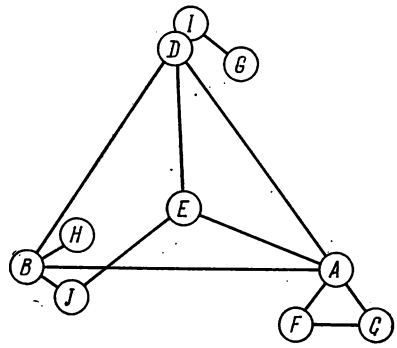
где D_{il} — кратчайшее расстояние между двумя точками СП, представляющими значения слов i и l, а d_{il} — алгебраическая разность между коор-

динатами i и ℓ на одном и том же измерении (факторе) j. Суммирование производится по всем факторам.

Применим данную формулу для вычисления расстояний между понятиями A и B из табл. 1.

$$D_{AB} = \sqrt{(-3-3)^2 + (3--3)^2 + (-3--3)^2} = \sqrt{72} \approx 8{,}49$$

По этой формуле можно вычислить расстояния между всеми парами понятий и построить на этой основе некоторую «концептуальную структуру» (авторы называют ее своего рода картой, «кусочком семантической географии»).



Работа по измерению расстояний между значениями слов в СП была продолжена Дж. Дженкинсом, У. Расселом и Дж. Сучи [108]. Они вычислили средние расстояния для 360 понятий в оценке 540 испытуемых на 20 шкалах, представлявших 6 факторов.

4. Закономерности взаимодействия значений. Рассмотрим словосочетания *умный человек*, *умная кошка*, *умный дурак*. На интуитивном уровне для нас более или менее ясно, что во всех этих случаях значения слов согласуются по-разному и в разной мере — больше в первом словосочетании, меньше в последнем. Можно предположить, что значения слов взаимодействуют в речи, причем 1) значение слова A влияет на значение слова B, и наоборот; 2) в результате этого влияния значения слов A и B модифицируются; 3) значение словосочетания AB в целом складывается из модифицированных значений слов A и B.

Перед нами возникают следующие вопросы: а) можем ли мы найти общие закономерности модификации значений сочетающихся слов и регулярно предсказывать меру модификации каждого значения в словосочетании, зная индексы значений, взятые изолированно? б) можем ли мы найти общее значение словосочетания, зная индексы значений входящих в него слов?

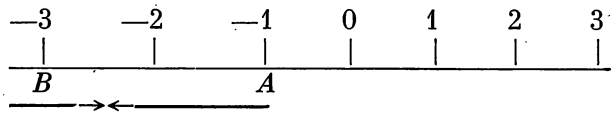
Оба эти вопроса получили интересную экспериментальную разработку в исследованиях Ч. Осгуда и его школы [50].

Как было отмечено выше, измерение значения методом СД дает две характеристики значения: направление (качество) и интенсивность (поляризацию). Когда два значения соединяются в словосочетании, то они 1) сдвигаются по направлению друг к другу (*умный* → *дурак*, *дурак* → *умный*), причем 2) степень сдвига обратно пропорциональна интенсивности значения: более поляризованное значение испытывает меньший сдвиг, и наоборот. Например, взаимодействие значений A

---

[50] Правда, в весьма ограниченных рамках; ср. критические замечания о методе СД в разделе III.

и B, занимающих точки —1 и —3 соответственно, может быть изображено следующим образом:



Сдвиг значения вычисляется по формуле

$$C_i = \frac{|p_j|}{|p_i| + |p_j|} \cdot P_i,$$

где C_i — сдвиг значения слова i, p_j — положение на шкале слова j, взаимодействующего с i, p_i — положение на шкале слова i, а P_i — разность $p_j - p_i$ (прямыми скобками обозначается модуль, т. е. некоторое число, взятое по абсолютной величине). Для нашего примера

$$C_A = \frac{|3|}{|1| + |3|} \cdot (-3 - -1) = -1{,}5; \; C_B = \frac{|1|}{|3| + |1|} \cdot (-1 - -3) = 0{,}5.$$

Точка соприкосновения стрелок интерпретируется как значение всего словосочетания. Оно вычисляется по формуле

$$d_m = \frac{|d_i|}{|d_i| + |d_j|} \cdot d_i + \frac{|d_j|}{|d_j| + |d_i|} \cdot d_j,$$

где d_m — значение всего словосочетания, $d_i = p_i$, $d_j = p_j$. Для нашего примера это дает

$$d_m = \frac{|1|}{|1| + |3|} \cdot -1 + \frac{|3|}{|1| + |3|} \cdot -3 = -2{,}5.$$

Авторы провели интересное экспериментальное исследование взаимодействия значений в словосочетаниях. Они взяли восемь существительных и восемь прилагательных и составили из них все возможные словосочетания. Каждое из 64 словосочетаний оценивалось испытуемыми на шкалах СД. Затем экспериментальные данные были сопоставлены с теоретическими величинами, вычисленными по второй формуле. Было обнаружено, что теоретические оценки незначительно отклоняются от эмпирических величин и являются вполне надежными [51].

5. Типологическое сопоставление концептуальных структур. Для того, чтобы выявить типологическое своеобразие двух концептуальных структур, необходимо показать, что они характеризуются хотя бы одним из следующих свойств: 1) сходные понятия получают в разных обществах разные оценки на одних и тех же факторах; 2) факторальная структура, действующая в одном обществе, отличается от факторальной структуры, действующей в другом обществе.

Различия первого рода были обнаружены в ряде экспериментов, проведенных сотрудниками Ч. Осгуда — Г. Кумата и У. Шрамом [116], которые изучали суждения студентов американских университетов (американцев, японцев и корейцев) о 30 политических понятиях, оценивавшихся

[51] С точки зрения Ч. Осгуда и его учеников, законы сочетания значений носят аддитивный характер. Можно предположить, что эта закономерность имеет ограниченную сферу действия — сочетания типа «прилагательное»+«существительное» и им подобные. Сочетания других типов могут подчиняться другим закономерностям. Сошлемся в этой связи на интересное экспериментальное исследование Н. Клиффа «Наречия как множители» [79], показавшего с помощью метода СД, что значения наречий и прилагательных сочетаются по мультипликативным законам. На этой основе Н. Клифф сумел оригинально объяснить причину того, что маловероятные, но вполне грамматичные сочетания типа «немного средний», «чрезвычайно обыкновенный» лишены смысла. Прилагательные «средний» и «обыкновенный» занимают на оценочной шкале точку, близкую к нулю, а умножение на нуль невозможно.

на 20 шкалах СД. Позже сходные результаты были получены Дж. Блоком [69], а также Г. Триандисом и Ч. Осгудом [169]. Дж. Блок обнаружил, например, что расстояние между понятиями «гнев» и «сочувствие» больше в английском языке, чем в норвежском. С другой стороны, американцы усматривают много общего в понятиях «горе» и «вина», в то время как норвежцам эти понятия представляются совершенно не связанными.

Различия второго рода в данных экспериментах не были вскрыты. Было обнаружено действие одних и тех же факторов. На этом основании Ч. Осгуд сделал вывод, что люди используют одни и те же основные измерения во всех своих осмысленных суждениях и что, следовательно, построенная им трехмерная модель СП является универсальной, т. е. не зависит ни от характера общества, ни от грамматического и лексического строя используемого в нем языка. Поэтому СД является единым стандартным методом для измерения значений любого языка и для межъязыкового сопоставления сходных понятий.

Новая серия экспериментов по типологическому сопоставлению концептуальных структур не подтвердила выводов Ч. Осгуда [117], [144], [161], [169]. Выяснилось, например, что в греческом языке в чистом виде представлен лишь фактор оценки. В нем же был обнаружен фактор, несопоставимый ни с какими другими известными факторами. Еще более разительные расхождения были обнаружены в ходе экспериментальных исследований факторальной структуры языка навахо [144], [166] сравнительно с факторальной структурой ряда других языков. Более или менее тождественными оказались лишь факторы оценки. В остальном факторальные структуры существенным образом отличались друг от друга.

Нам кажется, что единственный более или менее надежный, но вполне тривиальный результат этих исследований состоит в том, что оценочное измерение имеется во всех языках и что, по независящим от языка обстоятельствам, одни и те же понятия помещаются представителями разных обществ в разные места оценочной шкалы.

III. Замечания о теории СП

1. Замечания о факторальном анализе и модели СП. Мы не можем входить здесь в детали статистической техники факторального анализа [52]; для наших целей достаточно назвать некоторые ее математические недостатки и показать, как они могли отразиться на результатах Ч. Осгуда и его коллег.

Факторальный анализ не является полностью формализованной математической процедурой, поскольку алгебраическая матричная запись факторов и их геометрическая интерпретация в виде измерений многомерного пространства не связаны взаимно-однозначным соответствием. Предполагается, что в матрице представлена некоторая инвариантная факторальная структура, которой должно соответствовать единственное геометрическое решение. Фактически, однако, существует много различных геометрических решений, дающих разные факторальные структуры, и выбор одного из них диктуется не формальными, а чисто содержательными соображениями, т. е. является произвольным.

Эта сторона факторального анализа всегда была объектом серьезной критики, справедливо ставившей под сомнение ценность математической процедуры, которая опирается то на формальные, то на содержательные критерии [35], [37], [44], [74], [45]–[89]. Предпринятые позднее попытки

[52] Необходимые сведения о ней содержатся в обстоятельной книге Б. Фрюхтера [89], считающейся лучшим пособием в этой области, а также в двух статьях В. Д. Небылицына [44], [45]. В одной из них [44] имеется ясное изложение проблематики мультифакторного анализа и центроидного метода, которые особенно широко использовались в реферируемых работах.

формализовать все стадии факторального анализа не дали существенных результатов. Процедуру удалось формализовать лишь для тех случаев, когда число выделенных факторов не превышает двух. Наличие хотя бы трех факторов затрудняет принятие объективных решений. «В тех же случаях, когда анализом выделено 5, 10 и более факторов, проблема чрезвычайно усложняется. При ее решении может быть допущен значительный субъективизм, и в итоге два исследователя на одном и том же материале могут прийти к существенно различным выводам» [44].

Проанализируем с этой точки зрения материал, приводимый в книге «Измерение значения». В трех экспериментах, в которых Ч. Осгуд и его сотрудники обнаружили восемь различных факторов, эти факторы не оставались неизменными. Более или менее стабильными оказались лишь факторы оценки, силы и активности; факторы 4—8 в разных экспериментах выглядели по-разному: при изменении выборки шкал или испытуемых, или метода факторального анализа происходило перестроение факторальной структуры.

Во многих экспериментах даже трехмерное СП, образованное факторами оценки, силы и активности, не остается неизменным. Часто факторы силы и активности сливаются в один общий фактор динамизма, а иногда происходит слияние всех факторов в единый фактор оценки.

Факторальная структура меняется от анализа к анализу вопреки усилиям исследователя найти в таблице корреляций инвариантные соотношения, независимые от методов и условий факторизации. Это и естественно. Факторальный анализ, даже в наиболее совершенном виде, дает возможность выделить вполне надежно и строго лишь два фактора. При выделении большего числа факторов исследователю довольно часто и не всегда успешно приходится прибегать к интуиции и конструировать факторы на основе содержательных, а не формальных соображений.

В модели Ч. Осгуда вполне надежным является лишь фактор оценки, в меньшей мере — фактор динамизма.

Из сказанного следует, что факторальный анализ не является адекватным методом построения многомерного пространства. В таких обстоятельствах исследователь должен отказаться либо от идеи многомерности СП, либо от факториальной методики обработки экспериментального материала, которая ее дискредитирует. По всей видимости, разумнее сделать последнее, ибо идея многомерности СП представляется более привлекательной и продуктивной.

2. Замечания о методе СД и модели СП. Ч. Осгуд, Дж. Сучи и П. Танненбаум провели большую работу по проверке объективности, надежности, достоверности и чувствительности своего экспериментального инструмента и пришли к выводу, что СД в достаточной мере обладает всеми этими качествами [53]. С этой «официальной» авторской оценкой СД нельзя согласиться. Впрочем, и сами авторы во многих частях своей книги высказываются по поводу СД менее уверенно.

В теоретической части авторы утверждают, что СД измеряет значение, понимаемое как единство ситуативного (денотативного) и прагматического аспектов. В заключительной главе они признают, что СД эффективно измеряет лишь эмоциональную компоненту (прагматического) значения.

Сначала Ч. Осгуд и его коллеги склонны были считать, что суждения, индексируемые методом СД, обратимы. В этой связи они даже сравнивают свой инструмент с игрой в 20 вопросов. Однако в последней главе своей книги они признают, что суждения, закодированные языком СД, необратимы,[54] т. е. не поддаются декодировке. Действительно, если нам задано

---

[53] Аналогичная работа была проведена недавно У. Норманом [137].

[54] Позже факт необратимости суждений был экспериментально доказан в ряде остроумных опытов М. Маниса [126].

некоторое понятие, мы можем, пользуясь методом СД, найти соответствующую ему точку СП. Если, однако, нам задана некоторая точка СП, мы в принципе не можем реконструировать соответствующее ей понятие. Возьмем, например, точку СП, определяемую координатами «1 — *счастливый*», «1 — *твердый*», «2 — *быстрый*». Догадаться о том, что она соответствует понятию *отец*, можно только по чистой случайности.

Из сказанного можно сделать вывод, что, вопреки некоторым утверждениям авторов, СД не обладает достоверностью, так как он не ориентирован на тот объект (денотативное значение), который авторы определяют в теоретической части своей работы. Он далее, не удовлетворяет критерию чувствительности, потому что далеко не иденксирует различий, интуитивно ощущаемых говорящими. Наконец, в строгом смысле метод СД не является достаточно надежным, так как воспроизводимостью обладают далеко не все суждения.

Авторы «Измерения значения» склонны были считать, что постигшие их неудачи объясняются недостаточной широтой охвата материала. В связи с этим возникает следующий вопрос: может ли метод СД в его нынешнем виде, при условии более широкого охвата материала, привести к индексированию денотативного значения или он органически не приспособлен для этого? Анализ техники исследований методом СД показывает, что верно второе.

Коренной недостаток этой методики состоит в том, что она не обеспечивает однозначности и релевантности шкал. Все понятия «пропускаются» через все шкалы. Поэтому абсолютно все шкалы, используемые в экспериментах, имеют по крайней мере два разных значения: прямое и метафорическое. Когда на шкалах *гладкий — шероховатый*, *высокий — низкий* оцениваются понятия типа *столб* или *гриб*, шкалы принимаются испытуемыми буквально; когда те же шкалы используются для измерения значений таких слов, как *леди*, *грех*, шкалы могут быть поняты только метафорически. Большая часть шкал оказывается, таким образом, несущественной с точки зрения денотативного значения, которое остается поэтому невыявленным.

Поскольку подавляющее большинство суждений о понятиях, фиксируемых методом СД, носит метафорический характер, становится понятным отмеченное авторами в ряде экспериментов смещение всех шкал к оценочному измерению.

Итак, бедность осгудовской модели объясняется отнюдь не ограниченностью обследованного материала, но пороками, коренящимися в самой методике работы. Поскольку с помощью СД в его нынешнем виде можно выделить вполне надежно лишь одно измерение (оценки), СД не дает возможности строить многомерные пространства со сложной иерархической структурой.

СД должен быть серьезно усовершенствован еще в одном отношении: должна быть обеспечена не только однозначность шкал, но и однозначность вводимых понятий. Существующий метод ввода понятий не обеспечивает выбора нужного значения многозначного слова или нужного омонима. В результате одно и то же слово может быть воспринято разными испытуемыми в разных значениях, ср. *fire* ('огонь', 'пожар'), *engine* ('двигатель', 'локомотив'), *cop* ('«фараон»', 'верхушка', 'хохолок'. 'поимка'), *box* ('коробка', 'удар', 'самшит'), *trunk* ('чемодан', 'ствол'), *study* ('исследование', 'кабинет'), *mould* ('форма', 'плесень'), *stout* ('прочный', 'мужественный', 'тучный'), *plain* ('ясный', 'простой', 'некрасивый'). Во всех этих случаях весьма труден выбор так называемого «основного значения». Естественно, поэтому, что разные слова могут получить на той или иной шкале разные индексы не столько потому, что они оцениваются разными испытуемыми, сколько потому, что для оценки выбираются разные понятия.

3. Общая оценка семантической модели Ч. Осгуда. Смысл сделанных выше критических замечаний состоит не в том, чтобы дискредитировать колоссальную работу, проделанную группой Ч. Осгуда, а в том, чтобы показать, что исследователи не всегда осознавали стоявшие на их пути трудности и допускали поэтому серьезные просчеты. У. Вейнрейх метко и ярко определил значение поставленных экспериментов, сказав, что хотя исследователи не открыли нового пути в Индию, приобретенный ими опыт навигации может оказаться полезным [176].

Ценность описанных выше работ сводится к следующему:

1. Принципиально показано, что прагматическое значение может изучаться вполне объективно (экспериментально).

2. В некоторых отношениях экспериментальный метод Ч. Осгуда уже сейчас дает результаты, более объективные, надежные и богатые, чем результаты, получаемые интуитивно. Мы имеем в виду анализ оценочного компонента прагматического значения, исключительная роль которого была хорошо показана Л. Ельмслевом [102]. Операционное индексирование оценочного компонента существенным образом уточняет традиционные представления об этом аспекте прагматического значения и может быть с успехом использовано в прикладных целях.

3. Большое значение будет иметь впервые осуществленный в столь широких масштабах переход от двузначной логики к многозначной логике.

Е. ЗАКЛЮЧЕНИЕ

В заключение мы рассмотрим вопрос о соотношении синтаксических, сигнификативных, денотативных и прагматических значений. Он сводится к двум более частным вопросам: 1) как можно описать отношение множества значений различных типов друг к другу в данном языке и 2) как можно описать отношение множеств значений какого-либо одного типа в произвольно выбранной паре языков (изоформизм, взаимно-однозначное соответствие, одно-многозначное соответствие, много-многозначное соответствие).

Незавершенность современной лингвистической семантики не позволяет сколько-нибудь полно осветить эти вопросы. Полезно, однако, обсудить два принципиально противоположных подхода к решению указанных проблем, которые были намечены в этнолингвистике и математической логике. Эти два подхода нашли выражение в гипотезе Э. Сэпира — Б. Уорфа [54], [55], [56], развитой Г. Хойером [104], [105], [106], и в гипотезе Л. Витгентшейна — Б. Рассела [13], [48], [158], развитой Н. Хомским [77], [26].

Мы не будем обсуждать в данной работе детали той или другой гипотезы; они достаточно хорошо описаны в обширной литературе [8], [17],[68], [72], [81], [84], [94], [103], [118], [119], [120], [122], [139]. Мы не будем также оценивать этих гипотез: это было бы преждевременно. Для наших целей достаточно сопоставить основные идеи названных ученых.

Э. Сэпир и Б. Уорф утверждают, что у каждого языка есть своя неповторимая структура, свой «покрой». Каждый язык по-своему членит действительность и навязывает этот способ членения мира всем говорящим на нем людям. Язык оформляет мысль: люди, говорящие на разных языках, видят мир по-разному. При этом лингвистические модели связаны с культурно-социальными. Лингвистические различия, обязательные в данном языке, изоморфны различиям поведения, обязательным в данной культуре. Так, глагольная система языка навахо резко отличается от обычной в европейских языках системы обилием категорий, описывающих все аспекты движения и действия. Грамматические категории навахского глагола заставляют классифицировать в качестве разных объектов движение

одного тела, двух тел, более чем двух тел, а также движение тел, различных по форме и распределению в пространстве. Даже предметные понятия выражаются не прямо, но через глагольную основу, и поэтому предметы мыслятся не как таковые, а как связанные с определенным видом движения или действия. Г. Хойер говорит по этому поводу, что господствующей концепцией, пронизывающей весь строй языка навахо, является концепция мира, находящегося в движении «Параллели к этой семантической теме могут быть найдены в области культуры навахо, рассматриваемой в целом, — пишет далее Г. Хойер. — Даже в наши дни навахо являются по преимуществу бродячим, кочевым народом, перегоняющим свой скот с одного пастбища на другое. Мифы и легенды очень четко отражают этот мотив: боги и герои сказаний без устали путешествуют с одного святого места на другое» [104] [55].

В основе гипотезы Э. Сэпира — Б. Уорфа могут лежать два разных представления. 1) Внешний мир аморфен, лишен структуры, непрерывен. Язык дает нам возможность воспринимать его дискретно, организовать его. Поскольку нет единого эталона, унифицирующего языка, структуры языков различны. Поэтому мир членится разными языками и, следовательно, представляется говорящим на разных языках людям по-разному. 2) Мир определенным образом устроен, но в языке устройство мира отражается в сильно искаженном виде из-за обилия в нем случайных черт («шумов»). При этом шумы являются самой существенной чертой структуры языка. Без шумов язык перестает быть языком.

И в том и в другом случае общечеловеческая логика невозможна. Логика была бы совершенно другой, если бы ее написал не Аристотель, а представитель племени баганда. К этому выводу склонялся, в частности, Б. Уорф, не замечая, видимо, того, что тем самым ставится под сомнение возможность эффективного общения разноязычных обществ, многократно доказанная практикой межкультурных контактов.

Л. Витгенштейн и Б. Рассел выдвинули и обосновали другую точку зрения. Оставляя в стороне вопрос о том, является ли мир дискретным или непрерывным, они утверждают, что у ф а к т а (ситуации), имеется определенная (онтологическая) структура.

Факты, не разлагающиеся на другие факты, они называют атомарными (ср. «элементарные ситуации»). Например, сложный факт *Сократ был мудрый афинянин* состоит из двух атомарных: *Сократ был мудр* и *Сократ был афинянин*. Атомарный факт не содержит частей, являющихся фактами, но может быть разложен на составляющие его «простые» или «объекты» (ср. *Сократ* и *мудр*).

Объектам даются имена, из которых строится предложение. Предложение, утверждающее атомарный факт, называется атомарным предложением (аналог «ядерного предложения» в грамматике Н. Хомского). Атомарные предложения логически независимы друг от друга. Теория вывода (аналог «трансформационных правил» в грамматике Н. Хомского) относится не к атомарным предложениям, но к так называемым молекулярным предложениям (аналог «трансформов» в грамматике Н. Хомского), которые составляются из атомарных с помощью различных логических связок.

Далее Л. Витгенштейн и Б. Рассел выдвигают следующие ключевые положения своей теории: 1) имена и объекты находятся во взаимно-однозначном соответствии: «Мы не даем двух имен одной вещи или одного имени двум вещам» [48]; 2) синтаксическая структура атомарного предложения однозначно отражает онтологическую структуру атомарного факта (ср. [26]); 3) любое молекулярное предложение может быть однозначно

[55] О современном состоянии разработки гипотезы Э. Сэпира — Б. Уорфа см. [112]. Здесь же дается и весьма трезвая оценка гипотезы в свете последних экспериментальных данных.

определено в терминах первоначального множества атомарных предложений с помощью чисто формальных правил вывода [56].

Все сказанное относится к предложениям логически совершенного языка. В естественных языках имеются случайные и вредные свойства (шумы), искажающие фактически существующее в мире соотношение вещей. Эти случайные черты естественного языка связаны главным образом (но не исключительно) с теми обязательными грамматическими категориями, без выражения которых нельзя построить ни одного предложения (ср. сноску 34). Приведем еще такой пример [81]: вплоть до Э. Канта «существование» рассматривалось как предикат. Э. Кант впервые предположил, что «существование. . . не настоящий предикат» [81]. Ошибка происходила от того, что в естественных языках предложения типа *люди мыслят* и *люди существуют* имеют, как будто, одинаковую синтаксическую структуру. Это и заставляет предположить, что «мыслят» и «существуют» — элементы одного порядка. Раз «мыслят» — предикат, значит, и «существуют» — тоже предикат. В логическом языке ошибка такого рода исключена, так как высказывания, соответствующие приведенным выше предложениям, отличаются друг от друга синтаксически:

$$(x) : (x \text{ — человек}) \rightarrow (x \text{ мыслит}).$$

Для всякого x, если x — человек, то x мыслит.

$$(\exists x) : x \rightarrow \text{человек}.$$

Существует x, такой, что x — человек.

В отличие от Э. Сэпира и Б. Уорфа, Л. Витгенштейн и Б. Рассел не считают, что наличие этих случайных черт делает естественный язык алогичным. Они исходят из того, что логика едина для всех. Но единый логический язык, призванный однозначно отражать структуру мира, конструируется путем анализа слов и предложений естественного языка. Поэтому для любого предложения естественного языка можно в принципе восстановить его фундаментальную синтаксическую структуру, однозначно соотносящуюся с синтаксической структурой предложения на логическом языке и с логической структурой факта действительности.

Можно утверждать, что одним из весьма мощных средств установления скрытой логической структуры предложений естественного языка является их трансформационный анализ. Вернемся к приведенному выше примеру — фразам *люди мыслят* и *люди существуют*. Лишь на самом поверхностном уровне можно утверждать, что их синтаксическая структура одинакова. Они обнаруживают различные способности к синтаксическим трансформациям и поэтому должны считаться различными синтаксическими типами. В частности, тот факт, что «мыслить» — это предикат, а «существовать» — не предикат, проявляется в трансформации *люди мыслят* → *люди-мыслители*, которая невозможна для фразы *люди существуют* (ср. аналогичные примеры в части Б). Если семантический анализ слов и предложений будет дополнен трансформационным анализом, разработанным в структурной лингвистике, то это может открыть перспективный путь для решения важных лингвистических и логических проблем.

---

[56] Эта теория была и остается в логической литературе предметом оживленной полемики [17], [68], [72], [81]. Было показано, что она не имеет того универсального значения, которое первоначально склонен был приписывать ей Л. Витгенштейн [13]. Она не предусматривает многозначной и модальной логики и построена без учета того факта, что наши знания имеют до некоторой степени относительный и конвенциональный характер. Тем не менее, в определенных узких областях основные положения Л. Витгенштейна и Б. Рассела об атомарных и молекулярных предложениях сохранили свое значение для логики и лингвистики.

Литература

1. В. Н. Андреева. Лексикология современного французского языка. М., 1956.
2. Н. Д. Андреев. Об одной возможности моделирования семантики языка. ТДК.
3. Ю. Д. Апресян. Дистрибутивный анализ значений и структурные семантические поля. ЛС, № 5, 1962.
4. Ю. Д. Апресян. К вопросу о структурной лексикологии. ВЯ, № 3, 1962.
5. Ю. Д. Апресян. О понятиях и методах структурной лексикологии. ПСЛ.
6. И. В. Арнольд. Лексикология современного английского языка, Л., 1960.
7. Б. Б. Бирюков. Теория смысла Готлоба Фреге. «Применение логики в науке и технике», М., 1960.
8. М. Блэк. Лингвистическая относительность (теоретические воззрения Бенджамена Ли Уорфа), НЛ, I, М., 1960.
9. А. Л. Булаховский. Введение в языкознание, т. II, М., 1953.
10. В. В. Виноградов. Основные типы лексических значений слова. ВЯ, № 5, 1953.
11. О. С. Виноградова, А. Р. Лурия. Объективное исследование смысловых связей, ТКМП.
12. О. С. Виноградова, Н. А. Эйслер. Выявление систем словесных связей при регистрации сосудистых реакций, ВП, № 2, 1959.
13. Л. Витгенштейн. Логико-философский трактат. М., 1958.
14. В. Д. Волкова. О некоторых особенностях образования условных рефлексов на речевые раздражители у детей. «Физиологический журнал СССР», т. XXXIX, № 5, 1953.
15. И. Р. Гальперин, Е. Б. Черкасская. Лексикология современного английского языка, М., 1956.
16. Р. С. Гинзбург. Проблема значения слова. ИЯ, № 4, 1952.
17. И. С. Добронравов, Д. Г. Лахути, В. К. Финн. Комментарии к «Логико-философскому трактату» Л. Витгенштейна (см. 13).
18. А. В. Долгопольский. Изучение лексики с точки зрения трансформационно-переводного анализа плана содержания в языке, ЛС, № 5, 1962.
19. Л. Ельмслев. Пролегомены к теории языка. НЛ, I, М., 1960.
20. А. К. Жолковский. Рец. на сб. «Essays on and in Machine Translation by the Cambridge Language Research Unit», Cambridge, England, 1959. МППЛ, No 5, 1961.
21. А. К. Жолковский. Опыт составления списка слов для семантического анализа. ТДК.
22. А. К. Жолковский, Вяч. В. Иванов, Н. Н. Леонтьева, Ю. С. Мартемьянов, В. Ю. Розенцвейг, Ю. К. Щеглов. Дифференциально-семантические признаки слова. ДМК.
23. А. К. Жолковский, Н. Н. Леонтьева, Ю. С. Мартемьянов. О принципиальном использовании смысла при машинном переводе. «Машинный перевод. Труды института ТМ и ВТ АН СССР», вып. 2, 1961.
24. В. А. Звегинцев. Семасиология. М., 1957.
25. Вяч. В. Иванов. Теорема Геделя и лингвистические парадоксы, ТКМП.
26. Вяч. В. Иванов. Некоторые соображения о трансформационной грамматике. Тезисы докладов на конференции по структурной лингвистике, посвященной проблемам трансформационного метода. М., 1961.
27. А. Г. Иванов-Смоленский. Об изучении совместной работы первой и второй сигнальной систем мозговой коры. «Журнал высшей нервной деятельности», т. I, вып. 1, 1951.
28. Ф. И. Иващенко. Некоторые условия замены непосредственного раздражителя словом. ВП, № 3, 1957.
29. Р. Карнап. Значение и необходимость. М., 1959.
30. Н. И. Красногорский. К физиологии становления детской речи. Сб. «50 лет учения акад. И. П. Павлова об условных рефлексах», М.—Л., 1952.
31. Е. Курилович. Заметки о значении слова. ВЯ, № 3, 1955.
32. М. М. Ланглебен. Использование информационного языка для описания некоторых видов семантических связей между словами естественного языка. ДК.
33. Д. Г. Лахути, В. А. Успенский, В. К. Финн. Семантика. БСЭ, Изд. 2., т. 51.
34. К. А. Левковская. Лексикология немецкого языка. М., 1956.
35. А. Н. Леонтьев, А. Р. Лурия. XV Международный психологический конгресс, ВП, № 6, 1957.
36. Н. Н. Леонтьева. Модель синтеза русской фразы на основе семантической записи. ДК.
37. А. Р. Лурия. Впечатления о психофизиологических исследованиях в Англии. ВП, № 1, 1958.
38. М. Марушевская. О взаимодействии двух сигнальных систем в ориентировочных реакциях. ВП, № 1, 1957.

39. И. А. Мельчук. К вопросу о грамматическом в языке-посреднике. МППЛ, № 4, 1960.
40. И. А. Мельчук. Машинный перевод и лингвистика. ТМИЯ, гл. IV.
41. И. А. Мельчук. О некоторых типах языковых значений. ТМИЯ, гл. IV.
42. В. С. Мерлин. Динамика «переноса» условнорефлекторных связей из одной сигнальной системы в другую. ВП, № 2, 1957.
43. Е. А. Найда. Анализ значения и составления словарей. НЛ, II, М., 1962.
44. В. Д. Небылицын. Современное состояние факториального анализа. ВП, № 4, 1960.
45. В. Д. Небылицын. Факториальная структура интеллекта (обзор американских работ). ВП, № 2, 1961.
46. Е. В. Падучева. Возможности изучения языка методами теории информации, ТМИЯ, гл. VI.
47. В. И. Перебейнос. К вопросу об использовании структурных методов в лексикологии. ПСЛ.
48. Б. Рассел. Введение к «Логико-философскому трактату» Л. Витгенштейна (см. 13).
49. И. И. Ревзин, В. К. Финн. Рец. на кн. «The Measurement of Meaning» by Ch. E. Osgood, G. L. Suci, P. H. Tannenbaum 5, Urbana, 1957. ВЯ, № 4, 1959.
50. А. И. Смирницкий. Значение слова. ВЯ, № 2, 1955.
51. А. И. Смирницкий. Лексикология английского языка. М., 1956.
52. П. А. Соболева. Компонентный анализ значений на основе словообразовательного признака, ПСЛ.
53. В. Н. Топоров. К вопросу о возможности построения структурной лексикологии. ДМК.
54. Б. Л. Уорф. Отношение норм поведения и мышления к языку. НЛ, I, М., 1960.
55. Б. Л. Уорф. Наука и языкознание. НЛ, I, М., 1960.
56. Б. Л. Уорф. Лингвистика и логика. НЛ, I, М., 1960.
57. А. Черч. Введение в математическую логику. М., 1960. Введение.
58. А. Я. Шайкевич. Распределение слов в тексте и выделение семантических полей языка. ДМК.
59. С. К. Шаумян. Структурные методы изучения значений. ЛС № 5., 1962.
60. Л. А. Шварц. К вопросу о слове как условном раздражителе. «Бюллетень экспериментальной психологии и медицины», т. XXXVIII, вып. 12, 1954.
61. M. Aborn, H. Rubenstein. Perception of Contextually Dependent Word-Probabilities. AJP, vol. LXXI, No 2, 1958.
62. «Anthropology Today». An Encyclopedic Inventory. Prepared under the chairmanship of A. L. Kroeber. Chicago, 1957 (4-th ed.).
63. K. Baldinger. Die Semasiologie, Versuch eines Überblicks. Berlin, 1957.
64. Y. Bar-Hillel. Logical Syntax and Semantics. Lg., v. 30, No 2, 1954.
65. Y. Bar-Hillel, R. Carnap. Semantic Information. «The British Journal for the Philosophy of Science», vol. IV, No. 14, 1953.
66. H. Basilius. Neo-Humboldtian Ethnolinguistics, «Word», vol. 8, No 1, 1949.
67. E. Benveniste. Problèmes sémantiques de la reconstruction, «Word», vol. 10, № 2—3, 1954.
68. G. Bergmann. Two Criteria for an Ideal Language. PS, vol. 16, No 1, 1949.
69. J. Block. Studies in the Phenomenology of Emotions. JASP, vol. 54, 1957.
70. L. Bloomfield. Language. New York, 1933.
71. W. Bousfield, B. Cohen, G. Whitmarsh. Verbal Generalization: A Theoretical Rationale and an Experimental Technique, «Technical Reports», No 23, 1958.
72. M. Brodbeck. Coherence Theory Reconsidered: Professor Werkmeister on Semantics and the Nature of Empirical Laws. PS, vol. 16, No 1, 1949.
73. R. Carnap. Introduction to Semantics. Cambridge (Mass.), 1946.
74. J. B. Carroll. Рец. на кн. «The Measurement of Meaning» by Ch. E. Osgood, G. L. Suci, P. H. Tannenbaum, Urbana, 1957. Lg., vol. 35, No 1, 1959.
75. E. K. Charney. On the Semantical Interpretation of Linguistic Entities that Function Structurally. T., Paper 8.
76. N. Chomsky. Logical Syntax and Semantics. Their Linguistic Relevance. Lg., vol. 31, No 1, 1955.
77. N. Chomsky. Syntactic Structures. 's-Gravenhage, 1957.
78. N. Chomsky. Logical Structure in Language. «American Documentation», vol. VIII, No 4, 1957.
79. N. Cliff. Adverbs as Multipliers. PR. vol. 66, No I, 1959.
80. Contemporary Approaches to Cognition, Cambridge (Mass.), 1957.
81. I. Copilowish. Language Analysis and Metarhysical Inquiry. PS, vol. 16, No I, 1949.

82. A. Ellegård Рец. на кн. «Trends in Content Analysis», Urbana, 1959. — Lg., vol. 36, No 4, 1960.
83. «Essays on and in Machine Translation by the Cambridge Language Research Unit» Cambridge, England, 1959.
84. F. Fearing. An Examination of the Conceptions of Benjamen Whorf in the Light of the Theories of Perception and Cognition. LC.
85. J. Firth. The Technique of Semantics. «Papers in Linguistics 1934—1951». London, 1957.
86. J. Firth. Modes of Meaning. Ibid.
87. J. Flavell, E. Flavell. One Determinant of Judged Semantic and Associative Connection between Words. JEP, vol. 58, No 2, 1959.
88. Ch. Fries. Meaning and Linguistic Analysis. Lg., vol. 30, No 1, 1954.
89. B. Frûchter. Introduction to Factor Analysis, New York, Toronto, London, 1954.
90. E. Gibson. Рец. на кн. «Contemporary Approaches to Cognition», Cambridge (Mass.), 1957. — AJP, vol. LXXI, No 4, 1958.
91. W. Goodenough. Рец. на кн. «Language in Culture», Chicago 1954. — Lg., vol. 31, No 2, 1955.
92. W. Goodenough. Componential Analysis and the Study of Meaning. Lg., vol. 32, No 1, 1956.
93. J. Greenberg. The Logical Analysis of Kinship. PS, vol. 16., No 1, 1949.
94. J. Greenberg. Concerning Inferences from Linguistic to Non-Linguistic Data, LC.
95. P. Guiraud. La sémantique. Paris, 1954.
96. C. Haagen. Synonymy, Vividness, Familiarity, and Association. Value Ratings of 400 Pairs of Common Adjectives, JP, vol. 27, 1949.
97. Z. Harris. Componential Analysis of a Hebrew Paradigm. Lg., vol. 24, No 1, 1948.
98. Z. Harris. Methods in Structural Linguistics. Chicago, 1951.
99. Z. Harris. Distributional Structure. «Word», vol. 10, No 2-3, 1954.
100. Z. Harris. Co-occurrence and Transformation in Linguistic Structure. Lg., vol. 33, No 3, 1957.
101. F. Heider. Trends in Cognitive Theory. CAC.
102. L. Hjelmslev. Dans quelle mesure les significations des mots peuvent-elles être considérees comme formant une structure? «Reports for the Eighth International Congress of Linguistics», vol. 2, Oslo University Press, 1957.
103. Ch. Hockett. Chinese versus English: An Exploration of the Whorfian Thesis. LC.
104. H. Hoijer. Cultural Imlications of some Navaho Linguistic Categories, Lg., vol. 27, No 2, 1951.
105. H. Hoijer. The Sapir-Whorf Hypothesis, LC.
106. H. Hoijer. The Relation of Language to Culture. AT.
107. J. Jenkins, W. Russell, G. Suci. An Atlas of Semantic Profiles for 360 Words. AJP, vol. LXXI, No 4, 1958.
108. J. Jenkins, W. Russell, G. Suci. A Table of Distances for the Semantic Atlas. AJP, vol. LXXII, No 4, 1959.
109. K. Jones. Mechanized Semantic Classification, T, Paper 25.
110. M. Kay. The Relevance of Linguistics to Mechanical Translation. EMT.
111. E. Klima. Structure at the Lexical Level and its Implications for Transfer Grammar. T, Paper 13.
112. Ch. Kluckhohn. Notes on Some Anthropological Aspects of Communication. «American Anthrologist», vol. 63, No 5, 1961.
113. T. Knudsen, K. Sommerfelt. Principles of Uniligual Dictionary Definitions. «Reports for the Eighth International Congress of Linguists». Oslo University Press, 1957.
114. A. Kroeber. Classificatory Systems of Relationship. «The Nature of Culture», Chicago, 1952.
115. G. Kronasser. Handbuch der Semasiologie. Heidelberg, 1951.
116. H. Kumata, W. Schramm. A Pilot Study of Cross-Cultural Meaning. «Public Opinion Quarterly», col. 20, 1956.
117. W. Lambert, J. Hevelka, C. Crosby. The Influence of Language Acquisition Contexts on Biligualism, JASP, vol. 56, 1958.
118. H. Landar, S. Erwin, A. Horowith. Navaho Color Categories, Lg., vol. 36, No 3, part 1, 1960.
119. «Language in Culture», Chicago 1954.
120. D. Lee. Linguistic Reflection of Wintu Thought. IJAL, vol. X, No 2.
121. R. Lees. The Grammar of English Nominalizations, Chapter I. IJAL, vol. XXVI, No 3.
122. E. Lenneberg. Cognition in Ethnolinguistics. Lg., vol. 29, No 4, 1953.
123. L. Lipton, R. Blanton. The Semantic Differential and Mediated Generalization as Measures of Meaning. JEP, vol. 54, No 6, 1957.

124. F. Lounsbury. Field Methods and Techniques in Linguistics, AT.
125. F. Lounsbury. A Semantic Analysis of the Pawnee Kinship Usage. Lg., vol. 32, No 1, 1956.
126. M. Manis. Assessing Communication with the Semantic Differential. AJP, vol. LXXII, No 1, 1959.
127. M. Masterman. What is a Thesaurus. EMT.
128. M. Masterman. Fictitious Sentences in Language, EMT.
129. G. Matoré. Méthode en lexicologie. Domaine francaise. Paris, 1953.
130. C. Morris. Signs, Language and Behavior, New York, 1947.
131. C. Mosier. A Psychometric Study of Meaning. «Journal of Social Psychology», vol. 13, 1941.
132. R. Needham, A. Parker-Rhodes. This Question of Lattice Theory, EMT.
133. E. Nida. A System for the Description of Semantic Elements. «Word», vol. 7, No 1, 1951.
134. S. Newman. Semantic Problems in Grammatical Systems and Lexemes: A Search for Method. LC.
135. C. Noble. An Analysis of Meaning. PR, vol. 59, No 6, 1952.
136. C. Noble. The Role of Stimulus Meaning (m) in Serial Verbal Learning, JEP, vol. 43, No 6, 1952.
137. W. Norman. Stability Characteristics of the Semantic Differential. AJP, vol. LXXII, No 4, 1959.
138. C. Ogden, I. Richards. The Meaning of Meaning, New York, 1927.
139. D. Olmsted. Ethnolinguistics so Far. «Studies in Linguistics», Occasional Papers, No 2, 1950.
140. Ch. Osgood. The Nature and Measurement of Meaning. «Psychological Bulletin», vol. 49, 1952.
141. Ch. Osgood. Method and Theory in Experimantal Psychology, chapter VII. New. Lork, 1953.
142. Ch. Osgood. A Behavioristic Analysis of Perception and Language as Cognitive Phenomena. CAC.
143. Ch. Osgood. Semantic Space Revisited. «Word», vol. 15, No 1, 1959.
144. Ch. Osgood. The Cross-Cultural Generality of Visual-Verbal Synesthetic Tendencies. «Behavioral Science», vol. 5, 1960.
145. Ch. Osgood, T. Sebeok. Psycholinguistics. A Survey of Theory and Research Problems. Baltimore, 1954.
146. Ch. Osgood, G. Suci. Factor Analysis of Meaning. JEP, vol. 50, No 5, 1955.
147. Ch. Osgood, G. Suci, P. Tannenbaum. The Measurement of Meaning. Urbana, 1957.
148. W. Quine. On What there is, «From a Logical Point of View». Cambridge (Mass.), 1953.
149. W. Quine. Two Dogmas of Empiricism. Ibid.
150. W. Quine. The Problem of Meaning in Linguistics. Ibid.
151. W. Quine. Notes on the Theory of Reference. Ibid.
152. W. Quine. Word and Object, New York—London, 1960.
153. L. Rau. Variability in Response to Words: an Investigation of Stimulus Ambiguity. AJP, vol. LXXI, No 2, 1958.
154. H. Reichenbach. Elements of Symbolic Logic. New York, 1947.
155. R. Robins. A Problem in the Statement of Meaning. «Lingua», vol. III, No 2, 1952.
156. B. Ross, N. Levy. A Comparison of Adjectival Antonyms by Simple Card-Pattern Formation. JP, vol. 49, half I, 1960.
157. A. Rudskoger. Fair, Foul, Nice, Proper. A Study in Polysemy. Stockholm, 1952.
158. B. Russell. An Inquiry into Meaning and Truth, New York, 1940.
159. E. Sapir. Grading. A Study in Semantics. PS vol. II, No 2, 1944.
160. «Semantics». Encyclopedia Britannica.
161. C. Staats, A. Staats. Meaning Established by Classical Conditioning, JEP, vol. 54, No 1, 1957.
162. A. Staats, C. Staats. Meaning and *m*: Correlated but Separate. PR, vol. 66, No 2, 1959.
163. A. Staats, C. Staats, D. Biggs. Meaning of Verbal Stimuli by Conditioning. AJP, vol. LXXI, No 2, 1958.
164. A. Staats, C. Staats, W. Heard, L. Nims. Replication Report: Meaning Established by Classical Conditioning. JEP, vol. 57, No 1, 1959.
165. G. Stern. Meaning and Change of Meaning. Göteborg, 1931.
166. G. Suci. A Comparison of Semantic Structures in American Southwest Culture Groups. JASP, vol. 61, No 1, 1960.
167. A. Tarski. The Concept of Truth in Formalized Languages, «Logic, Semantics, Metamathematics», Papers from 1923 to 1938. Oxford, 1956.
168. A. Tarski. The Establishment of Scientific Semantics. Ibid.

169. H. Triandis, Ch. Osgood. A Comparative Factorial Analysis of Semantic Structures in Monolingual Greek and American College Students, JASP, vol. 57, No 1, 1958.
170. J. Trier. Der Deutche Wortschatz im Zinnbezirk des Verstandes. Heidelberg, 1931.
171. S. Ullmann. Words and Their Use. New York, 1951.
172. S. Ullmann. The Principles of Semantics, Glasgow, 1951; 2-d edition 1957.
173. S. Ullmann. Pièces de sémantique française, Berne, 1952.
174. S. Ullmann. Descriptive Semantics and Linguistic Typology. «Word», vol. 9, No 3, 1953.
175. U. Weinreich. Travels through Semantic Space, «Word», vol. 14, No 2—3, 1958.
176. U. Weinreich. A Rejoinder. «Word», vol. 15, No 1, 1959.
177. L. Weisgerber. Das Gesetz der Sprache. Heidelberg, 1951.
178. R. Wells. Meaning and Use. «Word», vol. 10, No 2—3, 1954.
179. R. Wells. To What Extent Can Meaning be Said to be Structured? «Reports for the Eighth International Congress of Linguists». Oslo University Press, 1957.
180. M. Wertheimer M., The Relation between the Sound of a Word and its Meaning. AJP. vol. LXXI, No 2, 1958.
181. C. Wimer, W. Lambert. The Differential Effects of Word and Object Stimuli on the Learning of Paired Associates. JEP, vol. 57, No 1, 1959.
182. C. Wordley. Establishing the Contextual Meaning of Words. EMT.
183. D. Worth. Transform Analysis of Russian Instrumental Constructions. «Word», vol. 14. No 2—3, 1958.

Сокращенные названия сборников и журналов

ВП Вопросы психологии
ВЯ Вопросы языкознания
ДК Доклады на конференции по обработке информации, машинному переводу и автоматическому чтению текстов, М., 1961
ДМК Тезисы докладов межвузовской конференции по применению структурных и статистических методов исследования словарного состава языка, М., 1961
ИЯ Иностранные языки в школе
ЛС Лексикографический сборник
МППЛ Машинный перевод и прикладная лингвистика.
НЛ Новое в лингвистике
ПСЛ Проблемы структурной лингвистики, М., 1962
ТДК Тезисы докладов на конференции по обработке информации, машинному переводу и автоматическому чтению текстов, М., 1961
ТКМП Тезисы конференции по машинному переводу, М., 1958
ТМИЯ О. С. Ахманова, И. А. Мельчук, Е. В. Падучева, Р. М. Фрумкина. О точных методах исследования языка. М., 1961
ТМК — Тезисы докладов межвузовской конференции по применению структурных и статистических методов исследования словарного состава языка. М., 1961.
AJP The American Journal of Psychology
AT Anthropology Today. An Encyclopedic Inventory Prapared under the Chairmanship of A. L. Kroeber. Chicago, 1957
CAC Contemporary Approaches to Cognition. Cambridge (Mass.), 1957
EMT Essays on and in Machine Translation by the Cambridge Language Research Unit. Cambridge, England, 1959
IJAL International Journal of American Linguistics
JASP Journal of Abnormal and Social Psychology
JEP Journal of Experimental Psychology
JP The Journal of Psychology
Lg Language
LC Language in Culture. Chicago, 1954
PR Psychological Review
PS Philosophy of Science
T The First International Teddington Conference on Machine Tsanslation and Applied Linguistics, National Physical Laboratory. Teddington, England, 1961

MODERN METHODS IN THE STUDY OF MEANING AND SOME PROBLEMS OF STRUCTURAL LINGUISTICS

Summary

This paper is a survey of modern methods in the study of structural, denotative, singnificative, and pragmatic meaning. Particular consideration is given to more or less rigorous methods of content analysis developed in the last decade in structural linguistics, semiotics, logical semantics, psycholinguisitcs and other related disciplines Among the most notable trends that are discussed and evaluated in the paper one may mention (1) analysis of meaning in terms of syntactic distinctive features, (2) componential analysis of meaning, (3) the semantic multiplier approach, (4) the thesauric approach of the Masterman group, (5) the psycho-physiological experiments of Luria and his colleagues, (6) the measurement of meaning as exemplified in Osgood's semantic space concept and some similar work.

II. ФОНОЛОГИЯ

А. Росетти

ЗНАЧЕНИЕ СПЕКТРАЛЬНОГО АНАЛИЗА ДЛЯ ФОНОЛОГИИ

И. Фодор [1] в своем исследовании о дифференциальных признаках в славянских языках подвергает строгой критике результаты спектрального анализа звуков. Результатом такого анализа, по его мнению, являются не фонемы, а звуковые типы, т. е. элементы, стоящие ниже фонем.

Разумеется, никакой анализ звуков с помощью технической аппаратуры не может обнаружить фонемы, т. е. такие элементы, которые выявляются лишь путем функционального анализа и которые поэтому находятся на ином классификационном уровне, нежели звуки речи.

Однако спектральный анализ позволяет проникнуть гораздо глубже в природу звуков речи, чем анализ звуков посредством той аппаратуры, которой пользовалась «классическая» фонетика (искусственное нёбо, кимограф и пр.). Спектральный анализ основан на акустических свойствах звуковой речи, которые в процессе общения важнее, так как именно они воспринимаются собеседниками.

Физиологический анализ звуков речи, основанный на артикулярных движениях произносительных органов, как известно, весьма приблизителен, так как органы речи находятся в постоянном движении. Опыт показывает, что можно произнести один и тот же звук разными способами. Наоборот, акустический анализ позволяет непосредственно проследить, как функционируют существенные и дифференциальные признаки фонем. Этот анализ дает возможность также установить дихотомическую классификацию звуков [2].

Обратимся к фактам.

Анализ звуков посредством «классической» аппаратуры дает возможность расклассифицировать звуки речи на основании произносительных критериев, например на звуки нёбные, губные и т. д.

Однако такая классификация слишком мало сообщает нам о существенных признаках каждого отдельного звука. Так, *k*, как и *g*, — нёбные со-

[1] I. Fodor. The role of the distinctive features and their use in the Slavonic languages. «Studia slavica». VII, Budapest, 1961, стр. 47.

[2] R. Jakobson. The phonemic and grammatical aspects of language in their interrelations. «Actes du VI-e Congrés international des linguistes», Paris, 1949, стр. 8: «Нужно анализировать две части знака: сторону физиологическую (и акустическую) и сторону психологическую», Z. S. Harris. Distributional structure. «Linguistics Today», New York, 1954, стр. 31: «Структура существует внутри говорящего субъекта»; R. Jakobson, M. Halle. Fundamentals of language. 's-Gravenhage, 1956, стр. 43 и сл.: «Дихотомический анализ фактов внутренне присущ структуре языка; двоичное противопоставление — это первая логическая операция ребенка»; B. Pottier. Problêmes de méthode en linguistique structurale. «Bollettino dell' Istituto di lingue estere», VI, Genova, 1961, стр. 7: «Двоичное противопоставление соответствует тенденции, присущей нашему мышлению».

гласные, и *k* отличается по признаку глухости от звонкого *g*. Здесь противопоставление действенно. Однако нельзя сказать того же о губных согласных. Здесь *b* и *m* — губные звонкие, а *p* и *f* — губные глухие. Таким образом, здесь противопоставление по звонкости и глухости отличает не отдельные звуки, а лишь пары звуков: звонкие (*b*, *m*) от глухих (*p*, *f*). Звуки *v* и *f* можно противопоставить *p* и *b* по несмычности. В отношении гласных, если считать румынские *e*, *a* и *o* звуками среднего подъема, а *i*, *l* и *u* — звуками верхнего подъема, можно провести классификацию, позволяющую отличить каждый гласный от всех остальных (*e* — нёбный звук среднего подъема, *i* — нёбный звук верхнего подъема, *ă* — неогубленный звук среднего подъема, *î* — неогубленный звук верхнего подъема, *o* — губной среднего подъема, *u* — губной верхнего подъема); однако термины этой классификации носят приблизительный характер [3].

Совсем иначе обстоит дело с классификацией звуков речи на основе спектрального анализа; здесь *k*, который наряду с *g* является компактным, противопоставляется звуку *g* по признаку напряженности (*k* — напряженный, *g* — ненапряженный; *b*, *p*, *m*, — это все диффузные звуки, но они противопоставлены друг другу по иным признакам: *b* — ненапряженный, *p* — напряженный, *m* — низкий. Гласный *o*, будучи компактным и низким звуком, противопоставляется *u* (тоже низкому, но диффузному), а *e* (компактный и высокий) противопоставляется *i* (тоже высокому, но диффузному) [4].

Таким образом, имеющиеся в нашем распоряжении средства анализа позволяют установить существенные признаки каждой фонемы, которые противопоставляют ее всем другим фонемам соответствующего языка.

Исходя из этих соображений, можно сделать оптимистическое заключение: непрерывный технический прогресс позволяет нам проникать все глубже в природу вещей.

Мы видим, что результатом спектрального анализа не являются фонемы, но этот анализ позволяет установить существенные и дифференциальные признаки фонем, полученные при помощи дихотомического анализа звуков.

Рассуждая так, мы исходим из принципа бинарности при анализе фонем. В случаях же, когда нужно учитывать три элемента, а не два [5], следует каждый раз решать, какой из признаков важнее и должен приниматься во внимание в первую очередь.

THE VALUE OF SPECTROGRAPH ANALYSIS IN PHONOLOGY

Summary

The apparatus which experimental phonetics makes use of cannot reach phonemes, nevertheless the improved ones, as for instance the Sona—graph, make a more subtile analysis of the spoken sounds, an analysis based on their acousitic qualities, and help us to estiblish the distinctive features of the phonemes. In this way we can establish a dichotomous classification of spoken sounds and this constitutes a progress in our research.

---

[3] См. B. Malmberg. La phonétique. Paris, 1954, стр. 65—66.

[4] См. R. Jakobson, G. M. Fant, M. Halle. Preliminaries to speech analysis. Massachusetts Inst. of Technology, 1955, стр. 43; E. Alarcos Llirach. Fonologia española. Madrid, 1961, стр. 82.

[5] A. Martinet. Economic des changements phonetiques. Berne, 1955, стр. 73 и сл. и M. Halle. In defense of the number two. «Studies presented to Joshua Whatmough», 's-Gravenhage, 1957, стр. 65—72.

Л. И. Богораз

О НЕЙТРАЛИЗАЦИИ И АРХИФОНЕМЕ В СВЯЗИ С СОГЛАСНЫМИ АРХИФОНЕМАМИ РУССКОГО ЯЗЫКА

1. Цель данной работы — привести в соответствие с современной лингвистической теорией некоторые факты русского литературного языка и одновременно уточнить соответствующие теоретические понятия. Речь идет о явлениях нейтрализации. Мы не будем останавливаться на критике теорий, не признающих понятий «нейтрализация» и «архифонема», а покажем лишь, что существующая трактовка этих понятий («не полностью идентифицируемые морфонемы» у М. Халле [4], стр. 328; «синкретизм» у Б. Мальмберга [2], стр. 356) приводит к противоречиям, не разрешимым в рамках реляционно-физических теорий.

Еще в 30-е годы Н. С. Трубецкой ввел в фонологию данные понятия и описал их в терминах своей системы. Приступая к исследованию нейтрализации, Трубецкой отмечает, что «нейтрализоваться могут только одномерные оппозиции» ([3], стр. 87). В дальнейшем развернулась полемика, главным образом, о том, принимать или не принимать эти понятия, имеют ли они теоретическое значение, какие явления в языке могут быть объяснены с их помощью. Что касается собственно определений Трубецкого, то они не были проанализированы в достаточной степени. Между тем их анализ может вскрыть следующее противоречие: нейтрализация — это утрата фонологически значимого признака парой фонем, противопоставленных друг другу только по данному признаку и отличающихся от всех остальных фонем системы совокупностью признаков, общих для этой пары, — именно таков смысл приведенного выше высказывания Трубецкого. Но каждая из фонем данной пары определяется лишь пучком дифференциальных признаков, устанавливаемых на основе фонологических противопоставлений с другими фонемами системы, и совершенно не обязательна столь симметричная система, чтобы ее фонемы составляли пары истинно одномерных оппозиций, т. е. были способны к нейтрализации. Остается либо расширить определение Трубецкого и признать, что нейтрализуемы не только одномерные оппозиции: тогда, сняв ограничение, мы лишаем понятие «нейтрализация» какой бы то ни было определенности и подводим под него все случаи аномальной дистрибуции; либо, последовательно применяя данное определение, мы настолько сузим круг явлений, объясняемых этим понятием, что оно фактически тоже лишится какой бы то ни было значимости, а исключенные таким способом явления все равно потребуют своего объяснения; либо, наконец, придется нарушить единство критерия и считаться как с фонологически существенными, так и с фонологически несущественными признаками. Именно так поступает сам Трубецкой: «. . . оппозицию *d*—*n*. . . можно рассматривать как одномерную, поскольку члены ее являются единственными звонкими дентальными

смычными (хотя ни звонкость, ни образование смычки не существенны для *n*, поскольку глухое и соответственно фрикативное *n* как особые фонемы в данной системе отсутствуют)» — [13], стр. 76. Однако, помимо неправомерности такого рассуждения самого по себе, оно подрывает самую основу понятия «нейтрализация», базирующегося именно на последовательном разграничении существенных и несущественных физических свойств.

Каждое из этих противоречий действительно неразрешимо на почве реляционно-физической теории. Такие понятия, как «нейтрализация» и «архифонема», получают строгое научное обоснование и объяснение в двухступенчатой теории фонологии, предложенной С. К. Шаумяном [5], [6]. Не претендуя на исчерпывающий анализ исследуемых явлений и их теоретических аналогов, мы попытаемся в данной работе уточнить определения нейтрализации и архифонемы, опираясь на факты, описанные уже в работах по русской фонетике и фонологии, и на двухступенчатую теорию фонологии.

Хотя предлагаемое в нашей работе описание фактов русского языка и не содержит новых наблюдений, мы считаем его необходимым здесь компонентом, так как любое описание связано с теоретическими предпосылками и задачами каждой конкретной работы; мы не могли, следовательно, ограничиться лишь ссылками на имеющиеся исследования по данному вопросу, а должны были предложить собственное изложение фактического материала, удовлетворяющее нашим задачам, и либо подкрепить выводы других авторов, либо — в отдельных случаях — дать иную трактовку описываемых явлений, согласующуюся с принятыми нами положениями.

2. Предварительные замечания.

2. 1. Нейтрализация речевых контрастов доступна непосредственному наблюдению. Анализируя речь с целью описания ее фонологической системы, исследователь может обнаружить, что контраст [1] между двумя реляционно-физическими элементами, отмеченный в ряде позиций, не наблюдается в некоторых позициях — нейтрализуется. Однако, к нейтрализации мы относим не каждый случай отсутствия контраста между двумя элементами системы в отдельных позициях (при его наличии в других положениях), а лишь те случаи, в которых имеем дело с элементами, а) связанными отношением взаимно-однозначного соответствия, б) в данной позиции представляющими случай свободного варьирования или дополнительной дистрибуции (эти положения мы принимаем, не подвергая их специальному обсуждению в нашей работе).

Нейтрализация речевых контрастов отражается на уровне конструктов как нейтрализация фонологических оппозиций.

2. 2. Для того, чтобы обнаружить нейтрализацию, полезно исследовать морфологические факты. Так, наличие минимальных пар на морфологическом уровне и исчезновение различия между ними в определенных условиях — это признак нейтрализации тех элементов, которыми данные пары различаются между собой. Однако морфологический критерий не является необходимым критерием фонологической нейтрализации и не может помочь в выяснении ее сущности.

2. 3. Нейтрализация вызывает изменения в системе фонологических оппозиций и требует конструирования новых систем, соответствующих данным условиям, данной позиции. Таким образом, в языке сосуществует несколько систем оппозиций, каждая из которых относится только к определенным условиям. Наиболее полная система относится к позициям, называемым сильными, все остальные — к слабым позициям.

---

[1] Термин «контраст» мы употребляем в трактовке С. К. Шаумяна ([5], стр. 17, 152), т. е. как обозначение отношения противопоставления на уровне наблюдения не только на синтагматической оси, но и на парадигматической.

2.4. Элементы систем, относящихся к слабым позициям, — архифонемы — отличаются от фонем сильных позиций именно тем, что относятся к другим системам. Обратившись к дифференциальным признакам фонем и архифонем, мы можем установить, что они состоят из разных наборов дифференторов, причем различие заключается в том,что у архифонемы на один дифферентор меньше, чем у соответствующих ей фонем (при одновременной нейтрализации двух оппозиций — на два, и т. д.). В то же время физические элементы, воплощающие архифонему, могут совпадать с тем или иным элементом, воплощающим фонему. И наоборот, физические элементы, воплощающие одну и ту же архифонему, в различных условиях могут быть совершенно различны. Иначе говоря, сходные физические элементы могут находиться в отношении воплощения к различным лингвистическим единицам — фонеме и архифонеме, а различные физические элементы — в отношении воплощения к одному и тому же члену фонологической системы, в нашем случае — к архифонеме.

2.5. Так как явление нейтрализации связано с условиями, в которых оно происходит, и вызывает изменение пучка дифференциальных признаков фонологических единиц в этих условиях, описание, очевидно, должно ответить на два вопроса: а) в каких позициях возникает нейтрализация данного противопоставления? б) что представляют собой фонологические единицы в этих позициях нейтрализации? Нам кажется, что удобно рассматривать каждую слабую для согласных позицию в отдельности, причем необходимо напомнить некоторые фонетические явления русского языка.

В современном русском литературном языке имеется несколько абсолютно сильных позиций для согласных, т. е. позиций, в которых различаются все согласные и к которым применима наиболее полная система оппозиций русских согласных фонем. Есть позиции, слабые по какому-нибудь одному признаку: звонкости/глухости, твердости/мягкости. В некоторых позициях совмещаются условия нейтрализации фонологического контраста по двум или более признакам.

Абсолютно сильными позициями для согласных являются положения перед гласными *a*, *o*, *u*, *i* (как под ударением, так и в безударном положении). В этих позициях различается 34 согласные фонемы: «m», «m'», «b», «b'», «p», «p'», «v», «v'», «f», «f'», «n», «n'», «d», «d'», «t», «t'», «z», «z'», «s», «s'», «c», «ž», «š», «ž:», «š:», «č», «g», «k», «x», «r», «r'», «l», «l'», «j». Положения в конце слова и в середине слова перед «е» и некоторыми согласными обладают меньшей различительной силой.

3. Нейтрализация контраста по признаку «звонкость/глухость».

3.1. В финальной позиции (на конце знаменательных слов) различаются следующие элементы: *m*, *m'*, *p*, *p'*, *f*, *f'*, *n*, *n'*, *t*, *t'* *s*, *s'*, *c*, *š*, *š:*, *č*, *k*, *x*, *r*, *r'*, *l*, *l'*, *j*: *kom*, *znăkon'*, *gr'ip* (*гриб* и *грипп*), *zip'*, *drof* (*дров* и *дроф*), *krof'*, *kon*, *kon'*, *plot* (*плод* и *плот*), *plot'*, *l'es* (*лез* и *лес*), *l'es'*, *ăt'ec*, *noš* (*нож* и *нош*), *doš:*, *plaš:*, *dač*, *mok* (*мог* и *мок*), *mox*, *v'er*, *v'er'*, *vol*, *vol'*, *voj*[2]. При сравнении с позицией максимального различения обнаруживаем, что в финале слова отсутствуют контрасты между *b* и *p*, *b'* и *p'*, *v* и *f*, *v'* и *f'*, *d* и *t*, *d'* и *t'*, *z* и *s*, *z'* и *s'* *ž* и *š*, *ž:* и *š:*, *g* и *k* — т. е. контрасты по признаку «звонкость/глухость». Таким образом, это противопоставление здесь нейтрализуется, а для единиц «*P*», «*P'*», «*F*», «*F'*«, «*T*», «*T'*», «*S*», «*S'*», «*Š*», «*Š:*», «*K*» «звонкость/глухость» перестает быть дифференциальным признаком, поскольку не отличает их от других единиц в той же позиции. Признавая, что фонемы можно идентифицировать по набору дифференциальных признаков, устанавливаемому на основе бинарных противопоставлений, мы

[2] Здесь и на стр. 156 примеры даны в фонемоидной транскрипции; ударение не обозначено, так как в фонемоидной и фонемной транскрипции ударные и безударные гласные обозначаются различными знаками.

должны считать эти единицы отличными как от звонких фонем «b», «b'», «v», «v'», «d», «d'», «z», «z'», «ž», «ž:», «g», так и от глухих «p», «p'», «f», «f'», «t», «t'», «s», «s'», «š», «š:», «k».

3. 2. Внутри слова, на стыке морфем, а также на стыке служебного слова со знаменательным тоже имеет место нейтрализация противопоставления «звонкость/глухость», однако особенности этого явления здесь отличны от первого случая. Для того, чтобы обнаружить нейтрализацию в данной позиции, следует принять во внимание контекст.

Перед звонкими *b*, *b'*, *d*, *d'*, *z*, *z'*, *ž*, *g* (т. е. перед всеми звонкими, кроме *v*, *v'*, *ž:*) элементы, воплощающие звонкие и глухие фонемы, совпадают в одинаковом звонком варианте; перед глухими *p*, *p'*, *f*, *f'*, *t*, *t'*, *s*, *s'*, *š*, *š:*, *k*, а также перед *č*, *c*, *x* (для которых глухость не является дифференциальным признаком) элементы, воплощающие звонкие и глухие фонемы, совпадают в глухом варианте: *pădb'il* — *ădb'il*, *ižzdal* — *zdal*, *ăbgăr'el*, *gžapədă*; *pătprav'il* — *ătprav'il*, *ištok*—*stok*, *ăpkrăt'il*, *ks'ev'irŭ*; *pətcœp'il* — *ătcœp'il*, *isxăd'il* — *sxăd'il*, *ăpčital*.

Таким образом, в указанных позициях по отношению к сильным позициям возникают не взаимно-однозначные позиционные преобразования, позволяющие объединить соответствующие друг другу элементы в один класс (см. [5], стр. 41—45), а много-однозначное соотношение, в котором двум элементам одного множества соответствует один элемент другого. Такие множества не эквивалентны, и их элементы не могут быть объединены в один класс.

Если же мы рассмотрим данную позицию независимо от сильных позиций, то обнаружим, что имеется взаимно-однозначное соответствие между звуковыми элементами перед первой группой согласных (пусть это будет P_1) и перед второй группой (P_2), причем элементу *d* на P_1 будет соответствовать *t* на P_2, *z* на P_1 — *s* на P_2, *g* на P_1 — *k* на P_2 и т. д. Физическое различие между *d* и *t*, *z* и *s*, *g* и *k* и т. п. (звонкость/глухость) обусловлено позицией и, следовательно, не является фонологически значимым. Значит, *d* и *t*, *z* и *s* и подобные им пары воплощают архифонемы «*T*», «*S*» и др., как и в первом случае, не обладающие дифференциальным признаком «звонкость/глухость».

Итак, противопоставление «звонкость/глухость» нейтрализуется в конце слова, а также перед согласными, которые сами являются членами этого противопоставления (кроме *v*, *v'* — перед ними различаются и звонкие, и глухие — и *ž:*, никогда не встречающимся после согласной), и перед *c*, *č*, *x*. Остальные возможные положения — перед всеми гласными и перед согласными *m*, *m'*, *v*, *v'*, *n*, *n'*, *r*, *r'*, *l*, *l'*, *j* — сильные для противопоставления «звонкость/глухость».

4. 1. Противопоставление «твердость/мягкость» в русском языке является взаимно-однозначным соотношением, следовательно, способно нейтрализоваться. Дифференциальный элемент «твердость/мягкость» обладает большим объемом, он представлен 12 парами фонем: «*m*»/«*m'*», «*b*»/«*b'*», «*p*»/«*p'*», «*v*»/«*v'*», «*f*»/«*f'*», «*n*»/«*n'*», «*d*»/«*d'*», «*t*»/«*t'*», «*z*»/«*z'*», «*s*»/«*s'*», «*r*»/«*r'*», «*l*»/«*l'*»; остальные 10 согласных фонем не обладают дифференциальным признаком «твердость/мягкость», хотя воплощающие их звуковые элементы либо постоянно твердые (*c*, *ž*, *š*), либо постоянно мягкие (*č*, *ž:*, *š:*, *j*), либо твердые или мягкие в зависимости от позиции (*g*, *k*, *x*).

Сильными в отношении твердости/мягкости оказываются положения перед всеми гласными, кроме *e(ĭ)*, а также конец слова, где различаются твердые и мягкие фонемы и архифонемы, возникшие в результате нейтрализации звонкости/глухости (назовем их архифонемами первой группы и будем обозначать прописной буквой с индексом 1′):

перед гласными — *mot*/*m'ot*, *bil*/*b'il*, *pil*/*p'il*, *val*/*v'al*, *grăfa*/*grăf'a*, *nit'*/*n'it'*, *vădа*/*văd'a*, *stuk*/*t'uk*, *grăza*/*grăz'a*, *sat*/*s'at'*, *luk*/*l'uk*, *rat*/*r'at'*;

в конце слова — *t'em/t'em'*, *kon/kon'*, *gol/gol'*, *žar/žar'*, *glüP$_{1'}$/gluP'$_{1'}$*, *kroF$_{1'}$/kroF'$_{1'}$*, *ploT$_{1'}$/ploT'$_{1'}$*, *l'eS$_{1'}$/l'eS'$_{1'}$*[3].

4.2. Рассмотрим позицию перед гласными *e* и *ĭ*. В русских корнях и в системе русского словоизменения в этой позиции нейтрализуется противопоставление «твердость/мягкость»: *ranə*, *ran'* — *ran'ĭ*; *păla* — *nəp̆ăl'e*, *păl'a* — *năpol'ĭ*; *klaT$_{1'}$*, *klaT'$_{1'}$* — *ăklad'ĭ*; *t'ma* — *t'm'e*, *ărba* — *ărb'e*, *tălpa* — *tălp'e*, *kăsa* — *kăs'e* и т. п. В приведенных примерах твердость/мягкость нейтрализуется перед падежной флексией *-e/-ĭ*; то же явление обнаруживается и внутри морфем: *zbor* — *b'ĭru*, *zăpor* — *zap'ĭr*.

Следует признать, как и в случаях 3.1. и 3.2, что здесь возникает не мягкая согласная фонема, а архифонема, для которой признак «твердость/мягкость» не является дифференторием. Довольно многочисленные заимствования с твердой согласной перед *e* (*mer*, *ser*, *fönetə* и т. п.) не нарушают этой закономерности: очевидно, что мы имеем здесь дело с сосуществующими фонологическими системами. В системе, применимой для заимствованных слов, правило нейтрализации твердости/мягкости перед *e* (как и некоторые другие правила) отсутствует, что, собственно, и позволяет выделить эту группу слов в особый отдел словаря. Н. С. Трубецкой отмечал, что «подобного рода слова..., попадая в обиходный язык, ... приспособляются к нормальным правилам сочетаемости фонем» ([3], стр. 284). Так, в слове *metr* (*мэтр* — учитель, наставник) нет нейтрализации — оно относится к особому отделу словаря; широко употребительное слово *m'etr* (*метр* — мера длины) подчиняется правилу о нейтрализации твердости/мягкости перед *e*. Можем ли мы считать пары слов *ser/s'er*, *mer/m'er*, *metr/m'etr* минимальными парами и сделать вывод о наличии оппозиции твердых и мягких согласных перед *e* в русском языке? Рассмотрим пару *ăd̂antω* (*о Данте*) — *ăd̂ant'ĭ* (*о Данте*). Без специального указания (в данном случае это фонемоидная транскрипция) мы не сумеем различить эти слова; они не составляют минимальной пары, так как относятся к различным, хотя и сосуществующим фонологическим системам и подчиняются различным фонологическим законам; из этого сопоставления мы также не можем сделать вывод об оппозиции «t»/«t'» перед *e(ĭ)* в русском языке, как из сопоставления слов различных языков[4].

4.3. В положении перед согласными обнаруживается неодинаковое «поведение» различных групп фонем, составляющих оппозицию «твердость/мягкость». Поэтому расчленим наше описание на отдельные случаи.

4.31. Губные фонемы «m», «m'», «b», «b'», «p», «p'», «v», «v'», «f», «f'» теряют признак «твердость/мягкость» перед любой согласной, причем в любом случае (кроме положения перед *j*) архифонема представлена твердым звуковым элементом: *siP'$_{1'}$* — *sωpnoj*, *sωpn'ak*; *droP'$_{1'}$* — *drobnəj*, *loP$_{1'}$* — *lobnəj*; *kroF'$_{1'}$* — *krăF$_{1'}$ca*; *zădumčĭvə*, *tərăpl'ivə*; *ziP$_{1'}$kə*, *riP$_{1'}$kə* (кое у кого сохранилось архаическое произношение *m'*, *b'*, *p'* ... перед мягким вариантом *g'*, *k'*, *x'*: *d'eF'$_{1'}$k'ĭ*, *proP'$_{1'}$k'ĭ*; но и в этом случае следует говорить о той же архифонеме, лишенной дифференциального признака «твердость/мягкость», только представленный иным, смягченным звуковым вариантом).

4.32. Дифференциальный признак «твердость/мягкость» сохраняет свое фонологическое значение для «*n*», «*n'*», «*d*», «*d'*», «*t*», «*t'*», «*z*», «*z'*», «*s*», «*s'*», «*r*», «*r'*» (и для соответствующих архифонем первой группы) перед *g*, *k*, *x* в любом варианте: *bankə/ban'kə*, *sankĭ/san'kĭ* (*Саньки*),

---

[3] Здесь и далее сохраняется фонемоидный принцип транскрипции, но в исследуемой позиции отмечаются архифонемы, о которых ранее уже шла речь.

[4] Положения перед безударными гласными мы здесь не касаемся, ограничиваясь пока более ясными случаями.

kaT$_1$ʻkə|kaT'$_1$ʻkə, *taS$_1$ʻkə|taS'$_1$ʻkə*, *varkə|var'kə*; *ăngar — d'ĭn'ga*, *pŭrga — s'ĭr'ga* (практически только перед *k*, перед *g* редко, перед *x* мы не обнаружили примеров).

Рассмотрим для этих же элементов позицию перед губными согласными. Перед твердыми губными, в одинаковом фонологическом контексте, встречаются и твердые, и мягкие зубные: *kărma — t'ăr'ma*, *ĭZ$_1$ʻba — r'ĭZ'$_1$ʻba*, *lis'va* (*Лысьва*, название реки) — *svat*, *smăla — t'ĭs'ma*, *pădmazəl — v'ed'mə*, *răzmil — kl'az'mə* (*Клязьма*). То же и перед мягкими губными (или губными архифонемами, воплощенными в мягком звуковом варианте): *kadm'ĭj — v'ed'm'ĭn*, *ărb'e — băr'b'e*, *ĭZ$_1$ʻb'e — r'ĭZ'$_1$ʻb'e*, *kărm'e — t'ăr'm'e*, *tm'in — d'ĭt'm'i*. Можно ли заключить, что твердость/мягкость зубных не зависит в данном случае от контекста, является самостоятельным фонологическим признаком? Будучи не обусловленной фонологическим контекстом, твердость/мягкость зависит здесь от характера соседствующих морфем: так, в конце приставки выступает, как правило, твердый звуковой элемент; конечная согласная корневой морфемы перед суффиксом *-ба* смягчена. Возможность свободного варьирования звуковых элементов — от твердого до полумягкого и мягкого — без риска нарушения взаимопонимания указывает на недифференцированность твердости/мягкости в данной позиции.

В фонологической литературе, даже у одного автора, нет единого мнения по этому вопросу. Так, Р. И. Аванесов в книге «Фонетика современного русского языка» [1] указывает, что «сильной позицией для твердости-мягкости согласных являются следующие положения: ... для переднеязычных — перед ... твердыми губными» (стр. 170—171), а в другом месте той же работы говорится: «Перед твердыми и мягкими губными выступают слабые по твердости-мягкости зубные» (стр. 181).

Перед *n*, *n'*, *d*, *d'*, *t*, *t'*, *z*, *z'*, *s*, *s'*, *r*, *r'*, *l*, *l'* твердость/мягкость зубных тоже не дифференцирована. Приведем несколько примеров (без транскрипции): *плен — пленный*, *пленник*, *конь — конный*, *конник*; *склад — складный*, *складни*, *медь — медный*, *медник*; *пот — потный*, *потник*, *плоть — плотный*, *плотнее*; *мороз — морозный*, *морознее*, *грязь — грязный*, *грязнее*; *двор — придворный*, *дворня*, *токарь — токарный*, *пекарь — пекарня*; *слюна — слюнтяй*, *лень — лентяй*; *талант — талантливый*, *прихоть — прихотливый*; *бант*, *бань* (глагол в повел. накл.) — *банте*, *баньте*; *пенс*, *пень* (глагол в повел. накл.) — *пенсе*, *пенься*; *рост*, *брось — росте*, *бросьте*.

В следующих примерах твердость/мягкость зубных перед согласными выступает как признак, зависящий от контекста: *кондор — кондёр*, *поддал — подделл*, *поднос — поднёс*, *подлый — подлец*; «n», «d», «t», «z», «s» перед твердыми согласными реализуются в твердых звуковых вариантах, а перед мягкими — в мягких и полумягких (а «r» почти всегда в твердом) вариантах.

То же можно сказать и о положении перед *c*, *č*, *ž*, *š*, *š:* (перед *ž:* — который используется в очень ограниченном числе морфем, как уже сказано, согласные вообще не встречаются), не участвующими в оппозиции «твердость/мягкость». Звуковые контрасты *n|n'*, *d|d'*, *t|t'*, *z|z'*, *s|s'*, *r|r'* нейтрализуются (или не различаются) перед этими согласными; возникающие в этих позициях архифонемы могут воплощаться в различных звуковых вариантах — перед *c*, *ž*, *š*, как правило, в твердых, перед *č*, *š:* — в мягких (кроме «*r*»), перед *j* — в твердых или мягких в зависимости от морфологического контекста.

И лишь оппозиция *l|l'* сохраняется перед любой согласной, кроме *j*.

4.4. Противопоставление «твердость/мягкость» в русском языке является не исконным, оно развилось в связи с изменениями в системе гласных фонем. Поэтому минимальные пары, различающиеся этим диф-

ференциальным признаком, часто оказываются различными грамматическими формами одного и того же корня, чередуются в вариантах одной морфемы (*графа — графя*, *черта — чертя* и т. п.). В это чередование в пределах вариантов одной морфемы включается и архифонема, для которой признак «твердость/мягкость» безразличен в силу определенных фонологических условий: *b'elәj — b'ĭl'oxәn'kәj — b'eL$_{1''}$ĭ n'kәj*. (Прописной буквой с индексом 1″ обозначена архифонема, нерелевантная по признаку «твердость/мягкость».) Такое чередование твердых и мягких фонем в вариантах одной морфемы, так же как и их нейтрализация, подчеркивает близкое родство этих фонологических единиц.

Если физические особенности звуковых элементов, воплощающих архифонемы, в случаях нейтрализации оппозиции «звонкость/глухость» зависят исключительно от фонологических условий, то выбор звукового элемента, воплощающего архифонему оппозиции «твердость/мягкость», часто определяется не фонологическим, а морфологическим контекстом. Архифонема здесь является не результатом нейтрализации в узком смысле слова, т. е. слияния двух фонологических единиц в одну, а элементом незавершенного процесса образования признака «твердость/мягкость».

Звуковые элементы в позициях, рассмотренных в 4.32, находятся в отношении свободной вариации (аналогичный случай указывает Базелл для аспирированных/неаспирированных в финальной позиции в датском языке — [7], стр. 27). Как говорилось выше, элементы, находящиеся в отношении свободной вариации, так же как и при дополнительной дистрибуции, не могут быть отнесены к двум различным классам, хотя бы в других окружениях они были противопоставлены. Следовательно, в позициях, о которых идет речь, противопоставление «твердость/мягкость» нейтрализуется; пучок дифференциальных признаков архифонем по сравнению с соответствующими фонемами сокращается на один признак, а именно тот, который отличает эти фонемы друг от друга.

5. Таблицы дифференциальных признаков согласных архифонем русского языка (табл. 1 и 2).

Сокращенные обозначения:

C (compact) — компактность
G (grav) — низкая тональность
N (nasal) — назальность
Cn (continuous) — непрерывность
L (liquid) — плавность
St (strident) — яркость
Ln (long) — долгота
S (soft) — мягкость
V (voiced) — звонкость

знаком + обозначено наличие данного признака, знаком — его отсутствие, 0 обозначает, что вопрос о наличии или отсутствии данного признака неприменим к рассматриваемой архифонеме.

5.1. Для случаев нейтрализации лишь одного противопоставления — «звонкость/глухость» или «твердость/мягкость» (с индексами соответственно 1′ и 1″) (табл. 1).

5.2. Есть позиции, в которых нейтрализуются оба противопоставления — и «звонкость/глухость», и «твердость/мягкость». Так, перед *k*(*g*, *x*), где нейтрализуется противопоставление «звонкость/глухость» и частично, для губных, «твердость/мягкость», возможен следующий состав фонологических единиц: «M_1», «P_2», «F_2», «*n*», «*n'*», «$T_{1'}$», «$T'_{1'}$», «$S_{1'}$», «$S'_{1'}$», «$Š_{1'}$», «c», «č», «$K_{1'}$», «*x*», «*r*», «*r'*», «*l*», «*l'*», «*j*» (индекс 2 отмечает архифонему, возникшую в результате нейтрализации обоих противопоставлений; «$Š{:}_{1'}$» отсутствует в списке, так как встречается только перед

Таблица 1

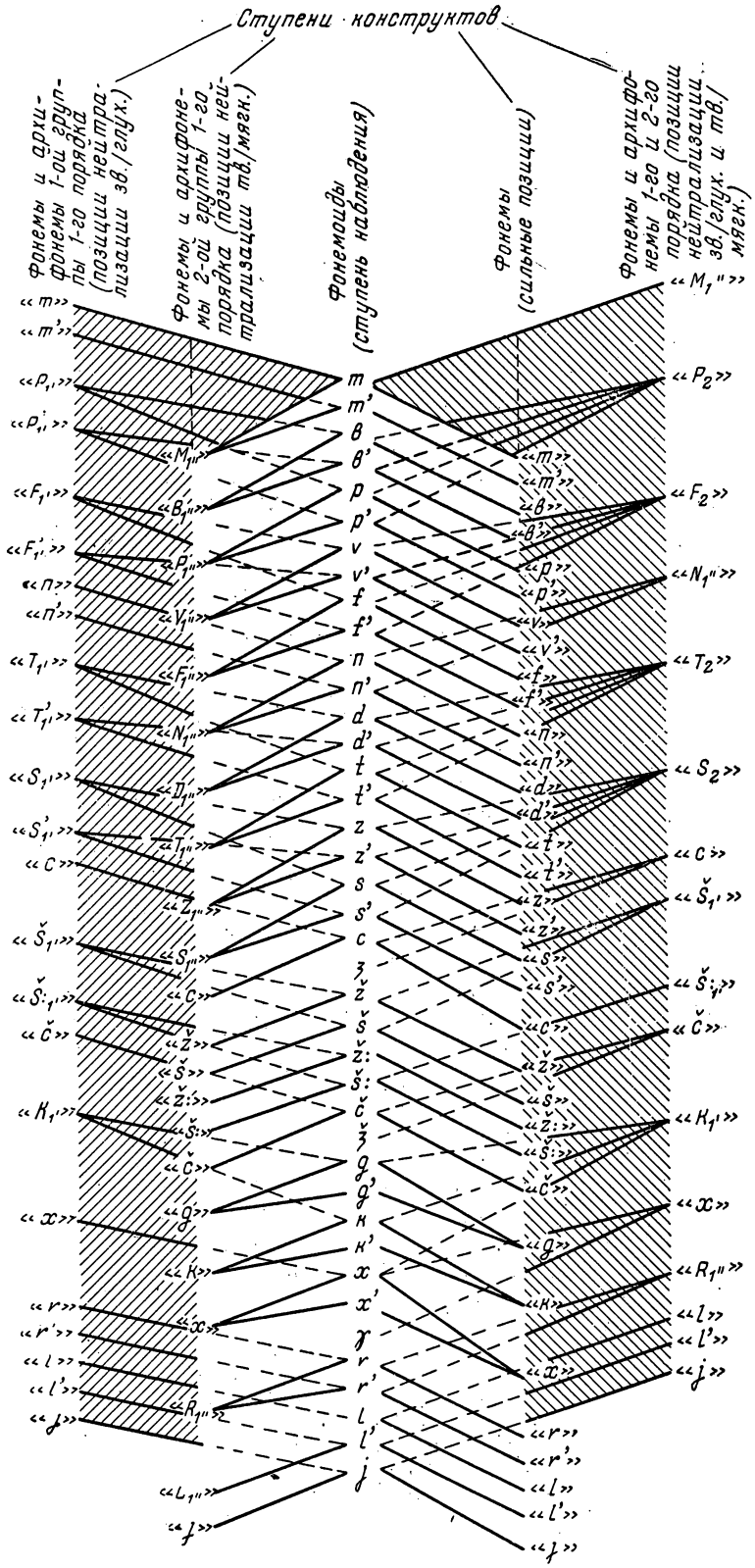

Ступени конструктов
Фонемы и архифонемы 1-ой группы 1-го порядка (позиции нейтрализации зв./глух.)
Фонемы и архифонемы 2-ой группы 1-го порядка (позиции нейтрализации тв./мягк.)
Фонемоиды (ступень наблюдения)
Фонемы (сильные позиции)
Фонемы и архифонемы 1-го и 2-го порядка (позиции нейтрализации зв./глух. и тв./мягк.)

Таблица 2

	C	G	N	Cn	L	St	Ln	S	V
$P_{1'}$	—	+	—	—	—	0	0	—	0
$P'_{1'}$	—	+	—	—	—	0	0	+	0
$F_{1'}$	—	+	0	+	—	0	0	—	0
$F'_{1'}$	—	+	0	+	—	0	0	+	0
$T_{1'}$	—	—	—	—	—	—	0	—	0
$T'_{1'}$	—	—	—	—	—	—	0	+	0
$S_{1'}$	—	—	0	+	—	0	0	—	0
$S'_{1'}$	+	—	0	+	—	0	0	0	0
$\check{S}_{1'}$	+	—	0	+	—	0	—	0	0
$K_{1'}$	+	+	0	—	—	0	0	0	0
$\check{S}:_{1'}$	+	—	0	+	—	0	+	0	0
$M_{1''}$	0	+	+	0	—	0	0	0	0
$B_{1''}$	—	+	—	—	—	0	0	0	+
$P_{1''}$	—	+	0	—	—	0	0	0	—
$V_{1''}$	—	+	0	+	—	0	0	0	+
$F_{1''}$	—	+	0	+	—	0	0	0	—
$N_{1''}$	0	—	+	0	—	0	0	0	0
$D_{1''}$	—	—	—	—	—	—	0	0	+
$T_{1''}$	—	—	0	—	—	—	0	0	—
$Z_{1''}$	—	—	0	+	—	0	0	0	+
$S_{1''}$	—	—	0	+	—	0	0	0	—
$R_{1''}$	0	0	0	—	+	0	0	0	0
$L_{1''}$	0	0	0	+	+	0	0	0	0

Таблица 3

	C	G	N	Cn	L	St	Ln	S	V
P_2	—	+	—	—	—	0	0	0	0
F_2	—	+	0	+	—	0	0	0	0
T_2	—	—	—	—	—	—	0	0	0
S_2	—	—	0	+	—	0	0	0	0

гласной и перед *n*): *suM*$_1$*kə*, *truP*$_2$*kə*, *broF*$_2$*kə*, *sank'ĭ*, *san'k'ĭ*, *r'eT*$_1$*kə*, *r'eT*$_1$*'kə*, *taS*$_1$*kə*, *taS'*$_1$*kə*, *noŠ*$_1$*kə*, *kl'ockə*, *nočkə*, *varkə*, *var'kə*, *kolkə*, *kol'kə*, *kojkə*. Перед остальными согласными, входящими в оппозицию «звонкость/глухость», и перед *c*, *č*, где утрачивают фонологическое значение оба дифференциальных элемента, тоже появляются архифонемы «P_2», «F_2», «T_2», «S_2»: *oP*$_2$*skəj* (Обский), *plăF*$_2$*ci*, *ăT*$_2$*d'el*, *r'ĭS*$_2$*ci*. Итак, в некоторых позициях появляются архифонемы, у которых двумя дифференциальными признаками меньше, чем у соответствующих фонем на абсолютно сильных позициях (табл. 3).

6. Заключение. Вернемся к тем противоречиям реляционно-физической теории, которые отражаются в фонологии русского языка

а) Если мы постулируем дифференциальные признаки фонем, исходя из контраста физических свойств звуковых сегментов и основываясь на оппозиции единиц более высокого уровня языка, то мы не можем считать финальный звуковой сегмент *t* идентичным звуковому сегменту *t* перед гласным: *t* в финальной позиции не противопоставлено *d*, следовательно, вопрос о том, звонкая фонема представлена здесь или глухая, неприменим для данного случая, в то время как фонема «t» перед гласным обладает дифферентором «звонкость/глухость». Реляционная точка зрения приводит к необходимости постулировать особый элемент фонологической системы — архифонему «T_1'» — и установить особое соотношение элементов фонологической оппозиции — нейтрализацию.

б) С той же реляционной точки зрения к фонеме «t» неприменим вопрос о назализованности или неназализованности, в то время как фонеме «d» должен быть приписан признак «неназализованность» (из противопоставления *d*—*n*); для «g» неприменим вопрос о прерывности/непрерывности, а в набор дифференторов фонемы «k» входит «непрерывность/прерывность» со знаком — (из противопоставления *k*/*x*). Таким образом, наборы дифференторов фонем «d» и «t», «g» и «k» отличаются друг от друга не только противоположными знаками в графе «звонкость/глухость», но и тем, что «d» обладает признаком «назализованность/неназализованность» (со знаком —), а «t» не обладает этим признаком; так же обстоит дело с фонемами «g» и «k». Но тогда основание для сравнения будет общим не только у «d» и «t», но и у «n», не только у «g» и «k», но и у «x» и в других подобных случаях — следовательно, каждая из этих оппозиций не является одномерной и, значит, не может нейтрализоваться.

Итак, возникло следующее противоречие: оппозиции «b» / «p», «d» / «t», «g» / «k» в определенных условиях нейтрализуются; но, не будучи одномерными, они, по утверждению Н. С. Трубецкого, не могут нейтрализоваться.

Это противоречие устраняется двухступенчатой теорией фонологии — разграничением двух ступеней абстракции и установлением строгих логических соотношений между ступенью наблюдения и ступенью конструктов. Действительно, оппозиции «b» / «p», «d» / «t», «g» / «k» не являются одномерными на ступени конструктов, где исследуются отношения между фонемами. Но понятие нейтрализации фонологических оппозиций соотносится с отсутствием определенных фонологических контрастов на ступени наблюдения, где исследователь устанавливает отношения между фонемоидами. Оппозиции (точнее, контрасты) фонемоидов *b* — *p*, *d* — *t*, *g* — *k* действительно одномерны, и в тех положениях, где эти контрасты исчезают, сохраняется различие между фонемоидом, воплощающим новую единицу системы — архифонему, и всеми прочими единицами системы. В теории на уровне конструктов это явление отражено в понятиях «нейтрализация» и «архифонема». Что касается классификации оппозиций по признаку одномерности или многомерности, то, играя важную роль на ступени наблюдения, этот признак утрачивает свое значение и становится несущественным на высшей ступени науки, где основой классификации служат более абстрактные отношения (см. [5], стр. 148—152).

Необходимо в связи со сказанным уточнить и понятие «архифонема». По определению Н. С. Трубецкого, архифонема — это «совокупность смыслоразличительных признаков, общих для двух фонем» ([3], стр. 87). Так как признак «непрерывность/прерывность» не является общим для фонем «g» и «k» (он отсутствует у фонемы «g»), из этого определения следовало бы, что «непрерывность/прерывность» — признак, нерелевантный и для архифонемы «K_1'». То же рассуждение при-

водит к признанию нерелевантным признака «назализованность/неназализованность» для архифонем «$P_{1'}$», «$T_{1'}$». Но в позиции нейтрализации звонкости/глухости сохраняется контраст между физическими элементами g и k, воплощающими архифонему «$k_{1'}$», и элементом, воплощающим фонему «х»; между звуковыми элементами d и t, воплощающими архифонему «$T_{1'}$», и элементом m, воплощающим фонему «m», и т. п. Исходя из этих противопоставлений, мы обязаны приписать архифонеме «$K_{1'}$» знак «—» в отношении непрерывности, архифонемам «$P_{1'}$» и «$T_{1'}$» знак «—» в отношении назализованности.

Итак, архифонема объединяет не признаки, общие для двух фонем, а все дифференторы каждой из них, за исключением того дифференторa, который, принадлежа каждой из фонем, имеет различные знаки и объединяет эти фонемы во взаимно-однозначное соответствие.

7. Схема, иллюстрирующая отношение систем архифонем и фонем русского языка к воплощающим их звуковым элементам, наглядно показывает, что одинаковые звуковые элементы воплощают различные фонологические единицы и что одним и тем же фонологическим единицам могут соответствовать различные звуковые сегменты.

Литература

1. Р. И. Аванесов. Фонетика современного русского литературного языка. М., 1956.
2. Б. Мальмберг. Проблема метода в синхронной фонетике. «Новое в лингвистике», II, М., 1962.
3. Н. С. Трубецкой. Основы фонологии. М., 1960.
4. М. Халле. Фонологическая система русского языка. «Новое в лингвистике», II, М., 1962.
5. С. К. Шаумян. Проблемы теоретической фонологии. М., 1962.
6. С. К. Шаумян. Двухступенчатая теория фонемы и дифференциальных элементов. ВЯ, 1960, № 5.
7. C. E. Bazell. Three conceptions of phonological neutralization. «For Roman Jacobson», Hague, 1956.

ON THE CONCEPTS OF «NEUTRALIZATION» AND «ARCHIPHONEME» IN CONNECTION WITH THE CONSONANT ARCHIPHONEMES OF RUSSIAN

Summary

The present paper is a fragment of a more copious investigation devoted to the positions of neutralization and to the consonant archiphonemes of modern literary Russian. The paper contains an examination of the most general and familiar cases of phonological neutralization: the neutralization of the voiced and voiceless, the hard and soft, and of both those features simultane onsly. New archiphonemes for which the feature of hardness/softness is not phonologically relevant are set up for the positions before m, b, p, v, f and the corresponding soft phonemes.

As a result of this analysis based upon S. K. Šaumjan's two-level theory of phonology more precise definitions of the concepts «archiphoneme» and «phonological neutralization» are proposed.

В. В. Шеворошкин

О СТРУКТУРЕ ЗВУКОВЫХ ЦЕПЕЙ

Можно классифицировать языки, исходя из количественного соотношения гласных и согласных в звуковых цепях. При этом следует помнить, что результаты подсчетов в значительной степени могут зависеть от того, все ли фонетические слоги фиксируются в записи текста именно как слоги, т. е. в виде (. . .*C*) *V* (*C*. . .). Лишь на основе записи или транскрипции, последовательно фиксирующей фонетические слоги, возможно выявление закономерностей сочетаемости элементов, относящихся к одному и тому же классу, — например, к классу согласных — и типологическая классификация этих сочетаний. Само то обстоятельство, что скопления элементов одного класса, обнаруживаемые в записи текста на данном языке, не поддаются достаточно четкой классификации, свидетельствует о нарушении в этой записи указанного принципа.

I. КЛАССЫ

Отношение количества гласных к количеству согласных в звуковых цепях различных языков мира колеблется между 2 и 0,5, не достигая, впрочем, этих чисел. Это отношение приближается к 2 в некоторых полинезийских языках (самоа, гавайском) и к 0,5 в германских языках (немецком, голландском и др.). Полинезийские и германские языки образуют тем самым два полюса, между которыми располагаются остальные языки мира, причем посредине оказываются такие языки, как суахили, японский или яванский, в которых данное отношение близко к единице. Языков, которые располагались бы за пределами обоих полюсов, по-видимому, не существует; мало того, в любом из языков, относящихся к одному из обоих крайних типов, можно обнаружить черты, как бы отрицающие достигнутую в процессе эволюции норму распределения *V* и *C* в звуковой цепи и свидетельствующие о том, что в дальнейшем изменение этой нормы пойдет (или уже идет) в обратном направлении.

Иногда полагают, что в полинезийских языках, в которых гласные обладают большей смысловой нагрузкой, чем согласные, существует тенденция к устранению согласных из звуковых цепей. В качестве примера приводят некоторые, в целом немногочисленные случаи, когда интервокальный согласный заменяется гласным — обычно *i* или *u* [1]. Уже простейшие подсчеты показывают, что в этих языках подавляющее большинство трехчленных сочетаний гласных — как в пределах слова, так

[1] Следует отметить, что оба слова — с сохранившимся и выпавшим согласным (согласными) — часто сосуществуют: ср. гавайск. *puamaru* || *puauau*. Что же касается достаточно длинных слов, состоящих сплошь из гласных, то они встречаются крайне редко (например, гавайск. *oiaio*), см. «Народы Австралии и Океании», М., 1956, стр. 565.

и в пределах фразы — содержат в качестве центрального члена именно гласные *i* или *u*, т. е. как раз те звуки, которые соответствуют неслоговым *i* и *u*, употребляющимся во многих языках (в частности, меланезийских) для устранения зияния и достижения более равновесного распределения гласных и согласных в звуковой цепи. Показательно в этом отношении и сосуществование пар типа *maori//maiori* (маори). В самоа, гавайском и некоторых других полинезийских языках все гласные характеризуются полным образованием, дифтонгов не существует, стечений согласных не допускается, исход слова или слога на согласный невозможен. Поэтому интервокальные *i* и *u* здесь полностью идентичны тем многочисленным *i* и *u*, которые встречаются в двухчленных группах гласных в конструкциях типа *VVC*-, *CVVC*-, *CVV*. Если упомянутую выше тенденцию к построению трехчленных групп гласных по типу *ViV*, *VuV* рассматривать как начало процесса, способного привести к устранению слогового качества интервокальных *i* и *u* — а именно об этом говорят перечисленные выше факты, — то нетрудно прийти к выводу, что процесс этот может в конечном итоге привести либо к образованию дифтонгов, либо к образованию двучленных консонантных элементов типа *i̯C*-, *u̯C*-, *Ci̯*-, *Cu̯*-, а также к появлению неслогового элемента в исходе слова.

О том, что наиболее богатые вокализмом полинезийские языки, по-видимому, развиваются в направлении меньшей насыщенности звуковых цепей гласными, свидетельствует, в частности, тот факт, что лексические пары типа *maori* // *maiori* (в определенной степени сопоставимые с явлениями типа индонез. *dia* $\geqslant$ [*dija*]) [2], наиболее характерны для языков, в которых преобладают группы *CV* (в самоа и подобных ему языках не менее часто встречается группа *CVC*). Возможность выделения в полинезийских языках двусложной группы *CVV*, как бы конкурирующей с односложным сочетанием *CV* (вокалическая насыщенность звуковой цепи в этих языках прежде всего зависит от того, насколько часто в них встречается группа *CVV*), сама по себе свидетельствует о том, что позиция гласного в звуковой цепи даже в языках, в которых гласные отличаются наибольшей автономией, в очень значительной степени зависит от структуры звуковой цепи в целом.

О том же свидетельствует и относительно небольшое количество трехчленных и более сложных комбинаций гласных даже в языках типа самоа: в этом языке на 1000 элементов звуковой цепи в среднем приходится всего 5—6 таких комбинаций в пределах слова и 19—21 комбинация на стыках слов. В целом гласные и согласные распределяются в звуковых цепях этих языков достаточно равномерно, чему, в частности, благоприятствует и то обстоятельство, что большинство слов начинается с согласного (исключение составляют лишь короткие служебные слова, часто содержащие звуки *i* и *u*).

---

[2] Ср. А. С. Теселкин, Н. Ф. Алиева. Индонезийский язык. М., 1960, стр. 12. Устранение зияния гласных, вполне объяснимое с точки зрения экономии речевых усилий, может быть охарактеризовано как проявление тенденции к восстановлению равновесия звуковой цепи, в особенности в языках, в которых $N_V : N_C \approx 1$. Кроме указанных неслоговых элементов, в индонезийском языке для устранения зияния используется гортанная смычка, разделяющая, например, два идентичных гласных (*keadaan* [ke$^{\text{ʔ}}$ada$^{\text{ʔ}}$an] — там же, стр. 12). Очевидно, нечто подобное имеет место и в полинезийских языках даже в тех случаях, когда предполагается, что два идентичных гласных ничем не разделены (если, разумеется, эти гласные действительно идентичны в фонетическом отношении). Два фонетически тождественных гласных, не разделенных смычкой, видимо, соседствовать не могут: *[aa] есть то же, что [ā], т. е. один долгий гласный, входящий в один слог. В этом отношении интересны примеры, приводимые Б. Галой в его содержательной, хотя и малооригинальной монографии (B. Hala. Slabika, její podstava a vývoj. Praha, 1956, стр. 50), в связи с полемикой с М. Дюраном, утверждавшим, что во франц. *Sahara* [saara] «rien ne sépare les deux *a*». Разумеется, при подобного рода исследованиях желательно основываться на данных экспериментальной фонетики.

Очевидно, «автономия» гласных в полинезийских языках весьма относительна и согласные в звуковых цепях представляют собой такую же опору для гласных, как и в любых других языках, с той только разницей, что один согласный нередко служит в полинезийских языках опорой для сочетания из двух гласных (в частых группах типа *CVV*). Дальнейшая вокализация звуковых цепей полинезийских языков могла бы идти по линии приближения структуры этих цепей к идеальной оптимальной схеме *CVVCVVCVV*. . ., однако этого не происходит. Но иного пути повышения вокалической насыщенности звуковых цепей, как это явствует из изложенного выше, не существует. Не существует и, по-видимому, не существовало языков с $N_V : N_C > 2$; в таких языках имелись бы многочисленные группы типа *VVV*. . ., что противоречит языковой реальности (пришлось бы предположить, что одной из обычных комбинаций является *CVVV*).

Уже из изложенного может последовать логический вывод, что в идеальном языке, звуковые цепи которого характеризуются отношением $N_V : N_C = 0{,}5$ (или $N_C : N_V = 2$), эти цепи должны отличаться равномерным следованием гласных и согласных. Было бы естественно предположить, что такие цепи строятся по схеме *CVCCVCCVC*. . . и состоят из сегментов типа *CVC*, т. е. из закрытых слогов.

Подсчеты показывают, что наибольшей консонантной насыщенностью звуковых цепей характеризуются не те языки, в которых основная смыслоразличительная нагрузка приходится на согласные и системы гласных отличаются крайней примитивностью (как это имеет место, например, в абазинском). Ни кавказские языки с их чрезвычайно разветвленными системами согласных, ни американские языки, которые часто приводятся как пример предельной насыщенности звуковых цепей согласными [3], на самом деле не являются полюсом консонантной насыщенности. Этот полюс образуют германские языки, в звуковых цепях которых отношение количества согласных к количеству гласных действительно приближается к 2, а сегменты цепей строятся в основном по схеме *CVC*. Тенденция к построению слогов — прежде всего, ударных — по схеме *CVC* или ее разновидностям и обусловленная этой тенденцией чрезвычайно высокая консонантная насыщенность звуковых цепей германских языков — в частности, немецкого — может быть проиллюстрирована следующими примерами. На четыре односложных французских слова типа *CV* в немецком приходится всего одно, причем односложные слова типа *CV* с кратким гласным отсутствуют вовсе [4]. Что касается открытых слогов немецкого языка с долгим гласным, то тенденция к их закрытию проявляется здесь в дифтонгизации долгих гласных, как и в структурно идентичном поведении трех сегментов: во-первых, групп *VC*, во-вторых, дифтонгов, т. е. групп типа «*V*+неслоговой элемент дифтонга», и в-третьих, долгих гласных, трактуемых в связи с такой постановкой вопроса, как «*V*+нечто» [5]. С другой стороны, имеется тенденция к прикрытию

---

[3] В американском языке кер д'ален, например, некоторые исследователи которого, по их собственному признанию, теряют дар речи при созерцании допустимых в нем комбинаций согласных (ср. R. A. R e i c h a r d. Coeur d'Allene. «Handbook of American Languages», Pt. 3. Hamburg—New York, 1933—1938, стр. 542), на каждые 100 гласных звуковой цепи приходится в среднем 150 согласных, т. е. столько же, сколько и в русском, тогда как в немецком на каждые 100 гласных приходится более 180 согласных (о закономерностях сочетаний согласных в языках мира, в частности в кер д'ален, см. ниже).

[4] Ср. J. Whatmough. Language: a modern synthesis. New York, 1956, стр. 216.

[5] См. W. F. T w o d d e l l. Рец. на кн.: F. P. M e n z e r a t h. Architektonik des deutschen Wortschatzes. —«Language», vol. XXX, 1954, стр. 593—597; об аналогичной интерпретации долготы в венгерском языке см.: Э. П. Х э м п. О долготе гласных в венгерском языке. ВЯ, 1960, № 1.

слога, которая выражается, в частности, в существовании твердого приступа [6], типологически сопоставимого с начальным индоевропейским ларингалом (отсутствие неприкрытых слогов), а также в трактовке простых интервокальных согласных в положении после краткого ударного гласного как относящихся к двум различным слогам — предшествующему и последующему. Действием все той же тенденции объясняется развитие в словах типа *Reihe*, *blaue* неслоговых [i̯] и [u̯], прикрывающих конечный гласный. Таким образом, нем. *kennen* ['kɛnən], имеющее в символической записи вид *CVCVC*, с точки зрения слоговой структуры может быть представлено в виде *CV*₵*VC*; что же касается тех редуцированных гласных, которые образуют открытый слог, то они подчас вообще исчезают (ср. *ich hab'* ['ɪç hap] из *ich habe*).

С другой стороны, сплошь и рядом обнаруживается действие контртенденции, или тенденции к уменьшению консонантной насыщенности звуковой цепи. В этой связи характерно разрушение той самой слоговой структуры *CVC*, которая служит одним из отличительных признаков языков германской группы. В нем. *Kinder* ([kɪndər] в традиционной транскрипции) последний слог представлен в наиболее распространенном современном произношении не в виде *CVC*, но в виде *CV*; [kɪndɑ] (что давно уже является нормой для английского языка). В этой же связи можно упомянуть отпадение конечного *n* в некоторых формах слов в голландском. В беглой речи очень часто изменяется слогораздел, в связи с чем нем. *hat er es* (в педантичной транскрипции [hat 'e:r 'ɛs]) произносится как [ha-te-res], что опять-таки ведет к увеличению числа открытых слогов.

Но даже если не учитывать подобных явлений, придется констатировать, что в немецком языке отсутствуют значительные скопления согласных в звуковых цепях, что прежде всего связано, естественно, с равномерным распределением согласных между гласными и с построением слогов по типу *CVC*. Так, в обследованном тексте на 1000 знаков транскрипции встретилось 29 групп типа *CCC* (считая и комбинации на стыках слов — эти группы в основном и встречались на стыках слов) и 3 группы типа *CCCC*, тогда как групп типа *CC* (без учета твердого приступа) на тот же текст пришлось 160 [7].

До сих пор речь шла о языках крайних типов — полинезийских с идеальной оптимальной структурой *CVVCVVCVVCVV* и германских с идеальной структурой *CVCCVCCVCCVC*. Даже в языках этих крайних типов встречаются значительные участки со структурой *CVCVCVCV*; в большинстве же языков мира этот последний тип структуры звуковых цепей преобладает. Значительные скопления элементов одного и того же класса (только гласных или только согласных) возможны во многих языках, хотя и далеко не во всех; но в целом такие скопления редки и встречаются, главным образом, на стыках слов. Очень важно подчеркнуть, что в языках, в которых допустимы однородные группы и гласных и согласных, всегда преобладают в частном отношении комбинации элементов какого-либо одного класса (только гласных или только согласных). Во всех языках, расположенных между двумя полюсами, в позиции между элементами одного и того же класса единичные элементы противоположного класса встречаются чаще, чем группы таких элементов. Другими словами, в подавляющем большинстве языков отдельные согласные

---

[6] На 30 слов французского языка типа *VC* в немецком приходится 26 слов того же типа — соотношение, не столь показательное, как в отношении слов типа *CV*; впрочем, если рассматривать твердый приступ в качестве особого согласного, то придется признать, что слова (и слоги) типа *VC* в немецком языке отсутствуют (возможны лишь слоги типа *CVC*).

[7] Подсчеты по кн.: О. Н. Никонова. Фонетика немецкого языка. М., 1958, стр. 217.

в интервокальном положении встречаются чаще, чем группы согласных, и, соответственно, в положении между согласными отдельные гласные встречаются чаще, чем группы гласных [8].

* * *

В соответствии со сказанным выше оказывается возможным наметить общий принцип поляризации гласных и согласных в фонетической записи текста на любом языке.

Рассмотрим следующие три случая:

1) Языки допускают стечения гласных, но не допускают стечений согласных. В таких языках трехчленные комбинации элементов типа *VCV* противопоставляются комбинациям типа *CVC*, *VVC*, *CVV*, *VVV* по центральному члену: в первом случае в качестве центрального элемента выступает *C*, в остальных случаях — *V*.

2) Языки допускают стечения согласных, но не допускают стечений гласных. Основанное на аналогичном принципе противопоставление будет иметь вид: *CVC* : *VCV*, *CCV*, *VCC*, *CCC*.

3) Языки не допускают стечений однородных элементов. В этом случае противопоставление будет иметь вид *CVC* : *VCV*.

Нетрудно заметить, что эти схемы противопоставлений имеют под собой реальные основания, с той только разницей, что в большинстве случаев бывают возможны более разнообразные сочетания (вплоть до *CCC*, *VCC*, *CCV*, *VCV* : *CVC*, *VVC*, *CVV*, *VVV*, например в русском языке). Однако в частотном отношении, как это следует из сказанного выше, по крайней мере один из комбинационных типов с данным центральным членом является доминирующим: так, в языках, в которых преобладают скопления согласных, из всех возможных комбинаций с центральным членом *V* доминирующей будет конструкция *CVC*, ибо комбинации этого типа встречаются гораздо чаще, чем комбинации *CVVC*; в языках, где $N_C : N_V \approx 1$, доминирующей трехчленной комбинацией с центральным членом *C*, естественно, будет *VCV*, а доминирующей комбинацией с центральным членом *V* — *CVC*.

Важно отметить, что в языках всегда имеется по крайней мере одна доминирующая комбинация, в которой центральный член противопоставлен периферийным элементам: *CVC*, *VCV*. Определение принадлежности элементов текста к одному из классов — к классу гласных или к классу согласных — в сущности сводится к выяснению того, какие именно трехчленные группы, идентичные в отношении центрального члена, противопоставляются друг другу по одному из указанных выше типов.

---

[8] В ряде языков, в особенности в языках с $N_C : N_V \approx 1$, например в корейском или якутском, имеется тенденция к устранению групп согласных, что ведет к построению более равновесной структуры звуковых цепей (о действующей в том же направлении тенденции к устранению зияния уже говорилось). Однако здесь, как и всюду, обнаруживается действие контртенденции, приводящей к образованию групп *CC* (или в более общем виде *CC*. . .) вследствие редукции гласного. Интересно, что в японском языке, в котором звуковые цепи строятся в основном по типу *CVCVCV*. . .., вследствие чего искажаются до неузнаваемости иностранные слова (ср. известные примеры Поливанова), также действует эта контртенденция, в связи с чем возникают начальные группы согласных типа *f*+*C* или *h*+*C* (где *f* и *h* относятся к подклассу *S*, см. ниже); отпадает конечный гласный (в связке *desu* и в глагольных окончаниях *-masu* и *-nu*), причем эти явления уже стали нормой (правильнее произнести [ftacu] 'два', чем [futacu], см. Н. И. Фельдман. Японский язык. М., 1960, стр. 18). Аналогичное противоречие обнаруживается и в тюркских языках, отличающихся весьма равномерным следованием гласных и согласных. Обе тенденции — к восстановлению равновесия звуковой цепи и к редукции гласных — вполне объяснимы с точки зрения принципа экономии речевых усилий.

Центральные элементы таких поляризованных групп и образуют объективно существующие полярные классы (класс гласных и класс согласных)[9]

Предположим, что нам необходимо выделить гласные и согласные из текста, зашифрованного следующим образом (известно, что каждое двузначное число соответствует определенному знаку текста, идентичного или очень близкого к фонетической записи):

01—02—03—04—05—06—03—06—07—04—08—09—10—11—02—07—06—08—10—03—07—12—02—13—04—07.

Выпишем столбцами все сочетания, идентичные в отношении центрального члена. Выписывать будем знак за знаком, так что, например, первые шесть знаков 01—02—03—04—05—06 войдут в состав не двух, а четырех столбцов: группа 01—02—03 войдет в столбец с центральным членом —02—, следующая группа 02—03—04 войдет в столбец с центральным членом —03—; в отдельные столбцы войдут и следующие две тройки: —03—04—05 и 04—05—06. Если текст ограничен, то можно использовать и начальные и конечные двойки: так, начальную двойку 01—02 можно включить в столбец с центральным членом 01 в виде 00—01—02, где 00 обозначает отсутствие знака (в данном случае, слева). Следуя этой методике, получим следующие столбцы:

1	2	3	4	5	6	7
(00—01—02)	01—02—03	02—03—04	03—04—05	04—05—06	05—06—03	06—07—04
	11—02—07	06—03—06	07—04—08		03—06—07	02—07—06
	12—02—13	10—03—07	13—04—07		07—06—08	03—07—12
						(04—07—00

8	9	10	11	12	13
04—08—09	08—09—10	09—10—11	10—11—02	07—12—02	02—13—04
06—08—10		08—10—03			

Рассмотрим вначале столбцы, содержащие наибольшее количество троек:

2	3	4	6	7
01—02—03	02—03—04	03—04—05	05—06—03	06—07—04
11—02—07	06—03—06	07—04—08	03—06—07	02—07—06
12—02—13	10—03—07	13—04—07	07—06—08	03—07—12

Поскольку нам необходимо прежде всего выявить группы типа CVC или VCV, естественно начать рассмотрение с комбинаций, крайние члены которых идентичны. Такие комбинации — если они вообще имеются в списке — представляют скорее всего именно группы CVC или VCV, так как в противном случае их пришлось бы отнести в группы CCC или

---

[9] Может возникнуть вопрос: почему именно анализ трехчленных сочетаний является универсальным методом выделения классов (а не анализ двучленных сочетаний, например)? Дело в том, что в двучленном сочетании принадлежность первого элемента к данному классу (C или V) еще ничего не говорит о природе второго элемента: так, в языках крайних типов однородные группы (CC и VV) встречаются чаще, чем разнородные (CV и VC), тогда как в языках промежуточных типов имеет место обратное. Что касается трехчленных сочетаний, то однородные трехчленные группы (CCC или VVV) в любом случае встречаются реже, чем разнородные группы, в частности группы CVC и VCV. Таким образом, если известно, что в данной трехчленной группе первый и третий элементы относятся к одному и тому же классу (C или V), то, вероятнее всего, центральный элемент относится к противоположному классу (V или C). Из приведенного выше определения частотности двучленных групп следует, что в языках центрального типа основным противопоставлением трехчленных групп будет противопоставление $CVC : VCV$, тогда как в языках крайних типов таким противопоставлением будет $VCV : VVC$, CVV, CVC и $CVC : VCC$, $CCV\ VCV$. Для языков, в которых $N_V : N_C$ или $N_C : N_V$ примерно соответствует 1,5, можно представить эти противопоставления в следующем виде: $VCV : CVC(VVC,\ CVV)$ и $CVC : VCV(CCV,\ VCC)$.

VVV, которые более редки. В нашем примере к таким комбинациям относится 06—03—06 (ст. 3). В столбце 6 дважды встречается 03 в качестве периферийного элемента в группах с центральным членом 06, в том же столбце «на периферии» дважды оказалось 07. С другой стороны, в столбце 7 дважды 06 оказывается в окрестности знака 07. В столбце 4 в окрестности знака 04 оказывается 03 и дважды 07. Этих данных достаточно, чтобы установить факт существования следующего противопоставления: 03, 07 : 04, 06, впрочем, не абсолютного, ибо ему противоречит структура следующих групп: 10—03—07 (ст. 3) и 03—07—12 (ст. 7). Из столбцов 2 и 3 следует, что 04 входит в тот же класс, что и 02, и что оба эти знака (02 и 04) противопоставлены знакам 03 и 07 (ст. 2: 01—02—03, 11—02—07; ст. 3: 02—03—04). Эти выводы подтверждаются и данными столбца 7. Противопоставление столбцов 2, 4 и 6 столбцам 3 и 7 (основанное на противопоставлении центральных элементов этих столбцов, а именно 02, 04, 06 : 03, 07) показательно, в частности, в том отношении, что в столбцах 2, 4 и 6 пока не обнаруживается противоречивых свидетельств, которые налицо в столбцах 3 и 7. Подтверждением того, что все комбинации столбцов 2, 4 и 6 действительно построены по одной из формул типа *x-y-x* и следовательно, образуют либо группы *CVC*, либо группы *VCV*, служат выводы, полученные из сопоставления столбцов 2, 4, 6; 3 и 7 с остальными столбцами. Так, в столбцах 4 и 6 в окрестности у знаков 04 и 06, входящих, согласно предварительному определению, в один и тот же класс, встречается знак 08; в столбце 8 в окрестности у этого знака встречаются не только знаки 04 и 06, но и знак 10, в окрестности которого встречаются знаки 03 и 08 в столбце 10; с другой стороны, в группе 04—08—09 обнаруживается противоречие в паре 08—09, что подтверждает отнесение знака 08 к классу 03, 07. . ., тогда как знак 10 относится к классу 02, 04, 06, что подтверждается свидетельством столбца 11, где знак 11 (входящий в окрестность знака 10 в столбце 10) имеет в качестве окрестности знаки 10 и 2; к тому же классу отнесем знаки 05 (04 и 06 в окрестности) и 13 (02 и 04 в окрестности). Нетрудно сделать окончательный вывод: знаки 02, 04, 06, 10 противопоставлены знакам 03, 05, 07, 08, 09, 11 и 13, причем регулярной структурой *x-y-x* характеризуются комбинации знаков, содержащие в качестве центрального члена знаки первой группы, а в качестве окрестности — знаки второй группы. Именно это обстоятельство заставляет отнести ко второй группе начальный знак 01 и знак 12, один раз встречающийся между знаками противоположных групп и расположенный после двух знаков второй группы. Больший контекст позволил бы объяснить эту нерегулярность: в зашифрованной латинской фразе *Basilus ut imperatum est facit*, которую мы только что проанализировали, последовательность —03—07—13—02— (*-st fa-*) приходится на стык слов. Отнесение знаков обеих выявленных групп к определенным классам (в данном случае, к классу *V* относятся знаки первой группы, а к классу *C* — знаки второй группы) основывается не на принципе сочетаемости или несочетаемости знаков одного и того же класса друг с другом, а на количественном различии в системе между *V* и *C*: первых всегда бывает меньше, чем вторых.

Подобным образом можно выделить два класса знаков как в текстах, записанных фонетической транскрипцией, так и в текстах, записанных близким к фонетической транскрипции буквенным письмом. (Разумеется, до тех пор, пока не будет поставлен вопрос о возможности передачи одного звука двумя знаками, сочетания вроде *ae*, *oe* в латинском будут трактоваться как группы гласных, а сочетания типа *mь* в русском — как группы «согласный+гласный» и т. д. Однако введение дополнительных критериев отбора позволяет решить положительно и этот вопрос)[10].

[10] Я не буду останавливаться здесь на этой проблеме, тем более, что в этом направлении ведутся детальные исследования (см. работы Б. В. Сухотина в сб. «Проб-

Как и в предыдущем случае, деление текста на слова не является обязательным. Рассмотрим знаковую цепь: *до сих пор мы исходили из того что сегментирование материала в нашей первоначальной записи было правильным.* Эта цепь может послужить основанием для составления следующих столбцов:

1	2	3	4	5	6	7	8	9	10	11	12	13
дос	оси	сих	ихп	хпо	орм	рмы	мыи	оди	или	изт	зто	ого
пор	исх	ыис	схо	йпе	иро	гме	был		ала	йза	что	егм
ход	осе	дил		апи	ери	ема	ным		аль		нти	
тог	иси	лии		опр	ерв				ыло		ате	
гоч		ииз			пра				иль			
тос		тир										
ров		ние										
вон		риа										
ной		пис										
лоп		сиб										
		вил										

14	15	16	17	18	19	20	21	22
очт	сег	ент	ова	ван	аше	ейп	льн	ибы
ача	мен	ани	авн	мат		ойз		
	ием	вна	рво	иал				
	тер	она	ави	лав				
	шей	ьно		наш				
	пер	ьны		нач				
				чал				
				зап				
				рав				

Сравним между собой, например, столбцы 2 и 3, содержащие полярные знаки *и* и *с* в качестве второго компонента. Анализ столбца 2 подтверждает разграничение *и* : *с* (в четырех сочетаниях четыре окрестных *и*) и позволяет сделать заключение об эквивалентности *и* — *о* (два сочетания с окрестным *о*). Столбец 1 (с центральным *о*) подтверждает противопоставление *о*, *и*, *а* : *л*, *ч*, *с* (одно *л*, одно *ч*, два *с*) и добавляет в группу *л*, *ч*, *с* еще и *в*, *ч*, *д*, *т*, *п*, *р*, *н* (по два знака в столбце 1). Большое количество этих знаков содержит столбец 18 с предварительно определенным *а* (по три *в*, *л*, *н*; два *ч*). Идя таким путем, можно с большой долей вероятности разграничить знаки *о*, *и*, *ы*, *е*, *а*, *ь* : *с*, *л*, *н*, *ч*, *в*, *м*, *д*, *б*, *ш*, *ч*. Знаки *п*, *р*, *т*, имеющие в своей окрестности знаки обеих групп, отнесем ко второй группе благодаря вескому свидетельству *о*, не имеющего в своей окрестности ни одного знака первой группы. В целом знаки первой группы имеют более однородную окрестность, допуская в нее из своей группы лишь знаки *и*, *ы*, *е*. На этом основании следует отнести знаки *х*, *й*, *з*, имеющие в своей окрестности по равному числу представителей обеих групп, во вторую группу. Остается определить, какая из групп содержит гласные и какая — согласные (в том, что перед нами буквенное письмо, сомнений не остается ввиду четкого разграничения двух типов знаков). Первая группа включает 6 знаков, вторая — 16, отсюда делается вывод о гласности звуков *о*, *и*, *ы*, *е*, *а*, *ь*, с ошибкой относительно *ь*, объяснимой исторически, и с частичной ошибкой относительно *е*. Нельзя также из небольшого текста сделать вывод о двояком характере знаков, передающих согласные. При достаточно большом контексте возможны более детализированные выводы.

лемы структурной лингвистики», М., 1962 и в настоящем сборнике). Важно отметить, что выделение классов и подклассов, разработанное Б. В. Сухотиным на основе совершенно иной методики, приводит в конце концов к результатам, сходным с изложенными здесь.

Поляризация классов знаков возможна и на основе текста в фонологической транскрипции, а также текста, записанного письмом, в котором не фиксируются редуцированные гласные, а слоговые сонанты не имеют специальных обозначений. Действительно, перевод текста типа *CVCCVC*. . . в текст типа *CCCVC*. . . не препятствует четкому выделению групп типа *CVC*, противопоставленных группам иных типов, поскольку в данном случае возрастание числа однородных скоплений *CC* (*C*. . .) не влечет за собой увеличения количества однородных групп типа *VV* (*V*. . .). Это значит, что выделение двух классов знаков возможно и на основе таких текстов, как калмыцкий или лидийский, в которых встречаются весьма разнообразные и многочисленные группы согласных букв. Калмыцкая и лидийская графика основана на фонологическом принципе: редуцированные гласные здесь не фиксируются (ср. аналогичный принцип английской транскрипции, в которой такие двусложные слова, как *function*, *pickled*, *rifles*, обозначаются как односложные: [fʌŋkʃn], [pɪkld], [raɪflz], а слоговые согласные не имеют специальных обозначений: ср. аналогичный принцип фонологической транскрипции санскрита [11]). Аналогичные приемы используются и в ряде других письменностей, обычно в ограниченной степени (например, в армянской графике; ср. также написание таких русских слов, как *остр*, *министр*, *подл* и т. д.).

Было бы желательно применить разработанную выше методику для дешифровки буквенных письменностей.

Как известно, до сих пор остается недешифрованным один из древних языков Малой Азии — карийский, о буквенном характере письменности которого весьма определенно высказался в 1950 г. Х. Боссерт [12]. Боссерт исходил из того обстоятельства, что в обнаруженной в 1949 г. Кавнской надписи на карийском языке на 240 знаков сохранившегося текста приходится всего 27 разных. Эта надпись может служить основанием для составления следующих столбцов [13]:

1	2	3	4	5	6	7	8	9
И	Н	Χ	⊕	F	C	M	X	Г
○ИН	ИНΧ	НΧ⊕	Χ⊕F	⊕FC	FCM	CMX	MXF	МГА
ΧИΨ	FНА	АΧИ	ϕ⊕F	XFC	VCF	FМГ	АХЕ	○ГМ
МИΨ	RНА	○ΧМ		AFR		ГМИ	АХР	МГ⊃Ϲ
FИА	○НР	МΧϕ		EF○		ΧМГ	ТХϕ	МГV
РИМ		АΧϵ		AFH		○МΧ	FXE	РГR
ϵИА		ΨΧV		AFϑ		ϕМА		RГϕ
VИ○				CFИ		RМГ		⩓ГR
МИМ				EF⋈		ϕМГ		МГϵ
МИФ				AFΨ		ИМ∇		ŤГϕ
FИР				○F∇		FМ○		○ГА
				○FM		СМ○		АГF
				○FX		ИМИ		ΨГА
				EFA		ΨМГ		
				VFP		ИМϵ		
				○FИ		ТМF		
				ГFV				
				MFA				

[11] В. В. Иванов, В. Н. Топоров. Санскрит. М., 1960, стр. 62 и табл. 3 (стр. 55); ср., однако, табл. 5 (стр. 58—59).

[12] См. замечания Х. Боссерта, приводимые в статье: F. Steinherr. Zu den neuen karischen Inschriften. «Jahrbuch für kleinasiatische Forschung», I, 1950/51, стр. 331 и сл.

[13] Факсимиле Кавнской надписи в уточненной редакции Деруа, дано в работе: L. Deroy. Les inscriptions cariennes de Carie. «L'antiquité classique», t. XXIV, f. 2, 1955, стр. 321.

10	11	12	13	14	15	16	17	18
A	Ψ	R	E	O	⋋	ϕ	V	⋈
ΓA⋋	ИΨΨ	FRΨ	ΨEF	FOΓ	Γ⋋R	F⋔O	ΨVC	F⋈A
ΨAF	ΨΨA	⋋RA	XEF	Ψ O⋋		⋋⋔X	⋈VΨ	T⋈V
XAF	RΨE	⋔RM	ϵEP	⋔OM		A⋔M	OVO	A⋈A
HAΨ	ИΨM	ΓRΓ	TEΨ	ΨO⍜		Ψ⋔R	ΓVT	
RAF	AΨO	ΓRϵ		⍜OP		Γ⋔M	⊖VИ	
ИAX	FΨ⋔	ΓRH		∇OV		A⋔P	⋔VF	
⋈A⋔	VΨO	ΔRM		VOF		ϵ⋔F	FVΨ	
MAF	⋔ΨP	ΨRT		ϵOF		Γ⋔P	⋋VF	
ΓAX	PΨO			MOΔ		Γ⋔⊕		
ϵA⋔	VΨ⋋			MOH		Δ⋔ϵ		
HA⋋	EΨR			ΨOF		X⋔P		
PA⋔	EΨΨ			ИO⍜		A⋔⊖		
ИAΔ	ΨΨM			⍜OΓ		⩓⋔V		
FA⋈				POΓ				
ΓAΓ				ΦOV				
ИAT								
FAΨ								

19	20	21	22	23	24	25	26	27
T	⍜	P	∇	⩓	ϵ	Δ	⊖	Φ
⋈TV	O⍜O	OPΓ	P∇O	∇⩓Γ	RϵA	OΔR	⋔⊖V	ИΦϵ
VTΓ		XPИ	F∇⩓		⋋ϵ⋔	AΔ⋔		ϵΦO
⋔TX		⋔P∇			ΓϵO			
RT⋔		HPA			PϵИ			
AT⋈		⋔Pϵ			ΦϵΦ			
ATM		ΨPΨ			PϵE			
		FPO			MϵИ			
		ИPϵ						

Нетрудно увидеть, что знаки A, R, E, O, ϕ, V почти не встречаются в окрестности друг у друга (некоторое исключение составляет группа A⋔, передающая, как это следует из более детального анализа карийских текстов, один гласный; то же может быть сказано и относительно OV), тогда как остальные знаки встречаются как в окрестности друг друга, так и регулярно в окрестности знаков первой группы. Тем самым знаки первой группы, образующие меньшинство, определяются как элементы класса гласных, остальные знаки — как элементы класса согласных [14]. Частые и бессистемные скопления согласных свидетельствуют, во-первых, о том, что в карийском тексте не фиксировались некоторые фонетические слоги, во-вторых, о том, что между двумя соседними согласными в ряде случаев проходил словораздел. Анализ таких карийских лексических пар, как ⋈MИA : ⋈⋔MИA, NFM⋔ : NVFM⋔, в текстах Египта может служить иллюстрацией спорадических консонантных написаний (характерно, что в текстах Египта, разделенных на слова, несколько сегментов встречается в сплошном консонантном написании [15]). Как известно, аналогичной структурой характеризовались и соседние с карийским ликийский и лидийский языки. Комбинаторный анализ

[14] Неясен характер редкого знака ⍜, встречающегося в текстах юго-восточной Карии только в позиции между двумя O. Знак ⊕, видимо, мог передавать в некоторых карийских текстах как согласный (очевидно, дентальный), так и гласный (очевидно, вариант [o]).

[15] Ср. комплекс ⊕FCM, дважды встречающийся в тексте из Кавна (ср. также XFCM).

Кавнской надписи позволяет выделить следующие сегменты, видимо, представляющие собой в большинстве случаев формы слов: МГА(Х/Ϫ), ТХ𐊫Р, ТГ𐊫Р, ∇ОVОF, ΨОF, ИР𐊴ЕР, МАFΨ𐊫, ХЕF(А)⋈А𐊫, 𐊴ИАΔ𐊫, ОГМИ(М), ⊖VИО♇ОГА; выделяются также следующие окончания и энклитики: -М, -Р, -F, Ψ, -Г, -Н, -(И)АТ, -(F-)С-М, -R-М-Г. Сравнение с другими карийскими текстами показывает, что графическая группа *CCVC* с исходным *C* = Р или М, видимо, представляет собой глагольную форму, включающую два фонетических слога и имевшую звуковую структуру *CVCVC*; сопоставление карийских энклитик и таких частых в лидийском энклитик, как *-k-m-λ-ad*, *-k-m-λ-t* и т. д., свидетельствует о том, что в карийском и лидийском языках имело место нанизывание друг на друга вспомогательных слов, включающих один согласный звук, сопровождаемый гласным призвуком. Наконец, анализ карийских слов в греческой передаче показывает, что в карийском языке имела место редукция гласных, возможно, еще более значительная, чем в соседних с карийским лидийском и ликийском языках [16]. В свете этих фактов получают предварительное объяснение необычные на первый взгляд комбинации карийских согласных букв в Кавнской надписи.

О том, что проведенная выше поляризация соответствует действительности, свидетельствуют другие карийские тексты, в частности тексты Египта [17], содержащие имена собственные, без труда идентифицируемые на основании предложенной транскрипции с карийскими именами, содержащимися в греческой передаче. Интересно, что каким бы путем мы ни шли к идентификации карийских букв, результат будет один и тот же: карийские буквы, так же как ликийские или лидийские, в основном соответствуют греческим как по форме, так и по звуковому значению. В пользу этого вывода говорит, в частности, уже тот факт, что в результате поляризации классов к гласным оказываются отнесенными в основном те карийские буквы, которые обозначают гласные звуки и в греческом, а к согласным — буквы, обозначающие в греческом (а также в лидийском и ликийском) согласные звуки. Поскольку карийская письменность, по-видимому, не предназначалась для целей тайнописи, такое положение вещей представляется вполне естественным [18].

II. ПОДКЛАССЫ

Гласные и согласные фонетической системы в целом соответствуют, таким образом, двум основным подклассам звуковой цепи, причем так называемым слоговым согласным в символической записи текста соответствует либо знак *V*, либо сочетание *VC*, *CV*. Очевидно, относительная равномерность гласных и согласных в звуковых цепях обусловлена, прежде всего, физиологическими возможностями речевого аппарата человека; в этом смысле предположение, что слова индоевропейского языка

[16] Ср. карийск. ⋈(𐊫)МИА (фонетически [mᵊsna], [mæsna]) и лидийск. Μασνης — ликийск. (в милийском диалекте) *masana*, хетто-лув. *Masna* 'бог'.

[17] См., например, тексты, содержащиеся в хрестоматии: J. F r i e d r i c h. Kleinasiatische Sprachdenkmäler. Berlin, 1932 (раздел «Karische Texte»). Кавнскому R здесь соответствует «коппа» (звуковое значение [i] которой было установлено Торпом на основании применения структурно-дистрибутивных методов уже около 60 лет тому назад); в текстах Египта, как и в некоторых текстах Карии, встречаются гласные, отсутствующие в Кавнском тексте: они обозначаются буквами, имеющими вид простого прямоугольника и прямоугольника с вертикальным штрихом внизу. Встречаются и некоторые другие буквы для согласных. Исследование показывает, что карийцы пользовались локальными алфавитами, причем число букв в каждом из них не превышало тридцати.

[18] Об идентификации карийских букв см.: В. В. Ш е в о р о ш к и н. Карийский вопрос. ВЯ, 1962, № 5; более детально эта проблема изложена в моей работе на английском языке «On Karian», подготовленной для журнала «Revue hittite et asianique» (Париж). — См. также мою заметку в Вя, 1963, № 3.

состояли из одних согласных (причем без всяких гласных призвуков), представляется абсурдным. С другой стороны, в ряде случаев приходится проводить резкую разницу между звуковой и фонемной цепью; так, если Глисон указывает на существование в американском языке белакула слов типа [tmk'mɬp], [sk'lxlxc], состоящих из одних согласных [19], то, разумеется, речь идет о фонемных цепях: соответствующие участки звуковых цепей характеризуются все тем же равномерным следованием элементов *C* и *V*, что и в любом другом языке.

Просмотрев исследования по классификации сочетаний согласных фонем, нетрудно прийти к выводу, что лежащие в основе этих сочетаний закономерности суть в сущности закономерности структуры звуковых цепей: лишь в тех случаях, когда фонемные и звуковые цепи оказываются идентичными в структурном отношении, можно говорить о том, что фонемные цепи отличаются четкой структурой (например, анализ сочетаний греческих согласных фонем, проведенный Н. С. Трубецким, как и разделы его работы, посвященные сочетаемости фонем и определению пограничных сигналов) [20]; в тех случаях, когда структура фонемных цепей оказывается резко отличной от структуры звуковых цепей, принципы построения фонемных цепей оказываются крайне сложными, так что все исследование сводится подчас к составлению пространственных списков возможных в данном языке сочетаний фонем [21].

Выявление различных позиционных типов согласных (подклассов) возможно как в рамках фонетического слога, так и на основе более обширных участков звуковых цепей (например, во французском и испанском языках). Естественно предположить, что подклассы соответствуют различным фонетическим типам согласных — смычным, сонантам, сибилянтам. В сущности так оно и есть, с той только разницей, что в одном и том же языке один и тот же согласный может входить в различные подклассы (в зависимости от его позиции относительно соседних с ним элементов сочетания), точно так же, как один и тот же звук может относиться к разным классам.

Так, согласный *v*, часто встречающийся в русском языке, наряду со звуками *f*, *s*, *z* и другими в позиции перед смычным или группой «смычный + сонант», например в словах *вдаль* (ср. *в тень*, *стал*, *сдал*), *вдруг* (ср. *в три*, *строг*, *здрав-*) входит в тот же подкласс, что и *f*, *s*, *z*, тогда как согласный *v* в словах *твой*, *с твоим* входит в тот же подкласс, что и звук *r* в словах *три*, *друг*; *в три*, *вдруг*. Нечто подобное имеет место и в отношении классов: действительно, какой бы незначительной с фонетической точки зрения ни была разница между *r̥* [22] и *r*, *i* и *i̯*, u и *u̯* (например, в санскрите; с точки зрения фонологии такой

---

[19] Г. Глисон. Введение в дескриптивную лингвистику. М., 1959, стр. 331; ср. обзор концепций относительно фонологии этого языка в статье: E. H. Haugen. The syllable in linguistic description. «For Roman Jacobson», the Hague, 1956, стр. 215 и сл.

[20] Сами по себе поиски закономерностей сочетаний согласных фонем недостаточно четко сообразуются с принципами современной фонологии: как известно, в фонологии не существует качественного различия между гласными и согласными фонемами; иное дело, если речь идет об элементах классов *V* и *C* звуковой цепи.

[21] Так, в работе «Об одном способе сочетаемости фонем в слоге (на материале классического тибетского языка)» Ю. К. Лекомцев указывает, что построение таблиц сочетаемости фонем в фонологическом слоге тибетского языка представляет собой «чрезвычайно сложную задачу» (см. «Краткие сообщения Института народов Азии», LVII, 1961, стр. 72). На резкое различие в структуре фонологических и фонетических слогов указывает, в частности, то обстоятельство, что частые и крайне разнообразные по структуре тибетские слоги с начальной группой *CCCC* не находят никакой аналогии в языках мира; фонетические (и фонологические) слоги с начальным *CCCC* встречаются крайне редко и имеют исключительно строгую структуру (ср. русск. *встр-*, *вздр-*, *вспл-*).

[22] Если это *r̥* не имеет никаких гласных призвуков и, следовательно, не может трактоваться как последовательная группа *ər* = *VC* или *rə* = *CV*.

разницы может и не быть[23]), в структурно-дистрибутивном отношении элементы *r*, *i*, *u* оказываются резко отличными от элементов *r*, *i*, *u*: первые входят в класс *V*, вторые — в класс *C*.

Рассмотрим оптимальную схему сочетаемости элементов, входящих в различные подклассы. В свернутом виде эта схема выглядит так:

$$(S)(T)(R) — V — (R)(S)(T).$$

Развернув ее, получим следующие типы сочетаний: *ST-*, *TR-*, *SR-STR-* перед гласным и *-RS*, *-ST*, *-RT*, *-RST* — после гласного. В подкласс *S* входят, как правило, спиранты, в подкласс *T* — смычные, а также аффрикаты, в подкласс *R* — сонанты. Данная схема в ее полном виде действует в языках с высокой консонантной насыщенностью звуковых цепей (например, в русском, в котором $N_C : N_V \approx 1{,}5$, или немецком, в котором $N_C : N_V \approx 1{,}8$)[24], в языках, в которых отношение $N_C : N_V$ лишь немного превышает единицу (например, в японском, дунсянском, яванском, индонезийском, финском, якутском, лакском и др.), используется один из редуцированных вариантов этой схемы (причем иногда бывает целесообразно вводить обозначение *C*: например, *-RC-* вместо *-RS-* и *-RT-* — единственно возможных типов сочетаний)[25].

* * *

Когда мы говорим, что во всех языках используется приведенная выше схема сочетаемости элементов или один из вариантов этой схемы, то имеются в виду либо все, либо большинство допустимых в данном языке сочетаний; если некоторое незначительное число сочетаний обладает особой структурой, то достаточно указать на этот факт, не прибегая к усложнению схемы (о ряде таких отступлений от общего правила см. ниже).

Иногда полагают, что те или иные языки (например, кавказские, некоторые американские, славянские) отличаются бессистемностью в отношении допустимых в них стечений согласных. Достаточно, однако, обследовать небольшой текст (в 300—500 знаков), чтобы прийти к противоположному выводу. Приводить примеры нет необходимости: подобный анализ, требующий всего нескольких минут, под силу каждому. Необходимо лишь принимать во внимание фонетические слоги.

Как следует из сказанного выше, схема $(S)(T)(R) — V — (R)(S)(T)$, как и некоторые ее варианты, может использоваться в пределах слога, тогда как другие варианты, обычно сильно редуцированные, могут характеризовать звуковую цепь в целом (так, в языках, в которых $N_C : N_V \approx 1$, например в японском, яванском, дунсянском, якутском, комбинации согласных отличаются весьма строгой структурой, в то время как сочетания согласных в пределах слога не допускаются). Наконец, существуют языки, в которых одни и те же типы комбинации согласных используются как в пределах слога, так и на стыке слогов (к таким языкам относится, например, лакский, отличающийся, как и многие другие кавказские языки, чрезвычайно строгими закономерностями структуры слова[26]).

---

[23] В. В. Иванов и В. Н. Топоров. Санскрит, стр. 52 и сл.

[24] Весьма близкий к указанной схеме структурный вариант, видимо, использовался в индоевропейском языке на определенной стадии его развития.

[25] В ряде случаев, в целях детализации исследования, возможно усложнение схемы; так, для русского языка можно вывести усложненную схему $(S_1)(S_2)(T)(R)—\ldots$, для готского $\ldots—(R_1)(R_2)(S)(T)$ и т. д.

[26] Ср. С. М. Хайдаков. Структурные типы лакских слов. «Вопросы изучения иберийско-кавказских языков», М., 1961, стр. 183—192. Отмеченная автором исключительная роль сонантов в структуре лакского слова (ср. Н. С. Трубецкой. Основы фонологии. М., 1960, стр. 308) указывает на использование в лак-

Разнообразие сегментов, в пределах которых комбинируются элементы класса C (отрезок звуковой цепи; фонетическое слово; слог), необязательно предполагает предварительное определение типа данного сегмента. Действительно, если предварительный анализ интервокальных групп согласных покажет, что эти группы характеризуются теми же структурными особенностями, что и сочетания согласных в анлауте, то само по себе это обстоятельство может свидетельствовать в пользу открытых слогов. С другой стороны, кажущаяся бессистемность интервокальных групп может свидетельствовать о наличии в структуре звуковой цепи многочисленных закрытых слогов. В этом случае целесообразно выявить вначале подкласс R, элементы которого характеризуются контактной позицией по отношению к вершине слога V. Одновременно с подклассом R выявляется и подкласс T (на основе противопоставления комбинаций $TRV : VRT$ или, в более общем виде, символических комбинаций $*TRT : RTR$). В последнюю очередь выявляется и «промежуточный» подкласс S[27].

Рассмотрим в качестве примера ситуацию, имеющую место в немецком языке: согласные *r*, *l*, *n*, *m*, *v* встречаются только по соседству с гласными, что само по себе может свидетельствовать об их вхождении в подкласс R (следует помнить, что речь идет о языке, в котором чаще, чем во многих других, встречаются интервокальные группы CCC). Любой отрывок текста покажет, что в немецком весьма часто встречаются группы *tr* + гласный и гласный + *rt*, а это может свидетельствовать о вхождении *t* либо в подкласс T, либо в подкласс R. На то, что *t* входит именно в подкласс T, указывают такие группы, как *∫tr-* (тип *STR-*), *-ft* (тип *-ST*) и т. д. Структурно сходны с *t* согласные *d*, *k*, *g*, *p*, *b*, *ts* (ср. группы *dr-*, *kr-*, *gr-*, *pr-*, *br-*; *kl-*, *gl-*; *-nt*, *-mt*; *tsv-*), что указывает на их вхождение в класс T; группы *∫n-*, *∫v-*, *∫r-*, *∫l-*, *fr-*, *fl-* образуют структуру *SR-*. Идя «изнутри» — от гласного к границе слога, — можно выявить закономерности структуры групп согласных в языке типа немецкого, которые окажутся сочетаниями, характеризующими слог.

Разумеется, если попытаться проанализировать все слова данного языка (в том числе и заимствования) с точки зрения их фонетической структуры, можно обнаружить множество групп согласных, строение которых не укладывается в схему $(S)(T)(R) — V — (R)(S)(T)$. Обычно такие группы составляют небольшой процент от общего числа сочетаний, встречающихся в потоке речи, причем часто в языке обнаруживается тенденция к устранению этих групп[28]. Тем не менее, некото-

---

ском языке следующего редуцированного варианта упомянутой выше схемы: $(C)V(R)(C)$. Аналогичное явление имеет место, по свидетельству автора, также и в других кавказских языках (там же, стр. 188).

[27] Таким путем, идя от гласного «вовне» слога, можно определить и границы слогов, морфем и слов. Внешне на такие границы, особенно на границу между словами, указывают — в особенности в языках с высокой консонантной насыщенностью звуковых цепей — хаотические скопления согласных (*CC*, *CCC*, а иногда и более сложные). Следует отметить, что мы не можем принимать здесь методику границ между языковыми единицами текста, разработанную З. Харрисом (Z. S. Harris. From phoneme to morpheme. «Language», 1955, vol. 31, № 2, стр. 190—223) и его последователями, ибо он исходит из предпосылки, что исследуемый язык заранее известен, мы же основываемся на том, что об исследуемом языке нам ничего не известно. На сходных предпосылках основаны и упомянутые в первой части настоящей работы чрезвычайно поучительные исследования Б. В. Сухотина.

[28] В этом случае также возможна аналогия с относительным распределением элементов V и C; допустимые в тех или иных языках однородные группы элементов представляют собой исключения из общего правила и встречаются весьма редко, причем вполне обычной является тенденция к устранению однородных комбинаций гласных путем перехода $V > C$, развитие V между двумя оказавшимися по соседству C и т. д.

рые из «нетипичных» комбинаций согласных встречаются в весьма распространенных словах. Рассмотрим некоторые из них—для примера обратимся к русскому языку, отличающемуся значительным разнообразием комбинаций согласных. Согласные *k* и *g* (*g′*), входящие в класс *T*, встречаются по соседству с элементами своего же класса *t*, *d* (*d′*), образуя необычные группы *TT* в словах *кто*, *где*. Эти слова употребляются в речи достаточно часто, чтобы развилась тенденция к переводу этих групп в более удобные антропофонические модели, а именно, в обычные группы *ST*-: показательно распространенное произношение [xto], [γd′e] > [d′e], нормальное произношение *что* как [ʃto]; устранение группы *TT* в *где* происходит еще и за счет отпадения начального согласного. В более редких словах, содержащих группу *TT*, оба начальных смычных могут сохраняться, что вполне соответствуют учению Е. Д. Поливанова о «минимальной трате»: ср. *бдеть*.

Отступления от нормы имеют место и в таких словах, как *мгла*, *рта*, специфику произношения которых иногда подчеркивают, вводя понятие побочного слога. Уже из приведенных примеров следует, что понятие «труднопроизносимой» группы согласных отнюдь не является фикцией: такие группы естественно возникают в ходе языковой эволюции и язык постепенно освобождается от них[29]. Что же касается конечных групп вроде *-стр*, то они с фонетической точки зрения представляют фикцию: слова *остр*, *быстр* — двусложные, а не односложные.

Иногда полагают, что необычные для европейских языков группы согласных (например, группа «носовой + смычный» в начале слога) представляют собой распространенное явление в языках других континентов. Однако в действительности «группа» *тр*-, например, соответствует в неевропейских языках либо образованию *тр*- (*тэр*-), отнюдь не представляя, таким образом, начального элемента слога[30], либо единому звуку, отсутствующему в языках Европы и передающемуся в графике и транскрипции посредством сочетаний типа *mb*[31]. Одному звуку соответствуют и написания *gb*, *kp*, *gbw*, *kbw*, *kpw* (ср. и.-е. g^w, k^w)[32] и т. д. Важно отметить, что принимаемые в современной транскрипции принципы обозначения необычных, «экзотических» (по отношению к языку составителя транскрипции) звуков сходны с теми принципами, которыми пользовались древние, когда им приходилось сталкиваться с необходимостью изобретения способа передачи тех звуков (=фонем) в их языках, обозначения которых отсутствовали в заимствованных ими алфавитах (ср. передача лабиовелярного посредством *крр* в ликийском — в позиции перед *a*[33]; характерно, что речь идет о языке, письменность

---

[29] Б. Гала в упомянутой выше работе на ряде примеров из чешского и словацкого языков доказывает неустойчивость «побочного слога» и перечисляет способы его устранения в языке: 1) отпадение начального согласного: чешск. (диал. и разг.) *sem* < *jsem*, *tuť* < *rtuť*; моравско-словацк. (диал.) *ňe* < *mňe*: ср. русск. разг. [ď′e] < *zde*; 2) развитие нормального слога на месте побочного: чешск. *sedem*, старочешск. *listivě*; 3) метатеза: чешск. (диал.) *hlář*, *truť*, *smíce*; 4) протеза: словацк. *eržať*, *ortút′*, *omša*; ср. русск. (диал.) *аржаной*; 5) переразложение слога: моравско-словацк. *pom-ně* (=*po mně*) (ср. русск. разг.-простореч. *мине* < *мне*); чешск. *něl-ži*, *naj-mu* и т. д.; 6) изменение качества согласного: чешск. *včela* > > *fčela*, *ptát* > *ftát* (ср. выше относительно русск. *кто* > разг. *хто*). Подобного рода примеры как нельзя лучше свидетельствуют против положений А. Исаченко, развитых в его известной статье «Опыт типологического анализа славянских языков» («Новое в лингвистике», вып. III, 1963).

[30] Emmi Dolezal Baule: Sprache eines Negervolkes an der Elfenbeinküste. «Die Sprache», 1958, Bd IV, стр. 181—182.

[31] Ср., например: Е. Н. Мячина. Язык суахили. М., 1960, стр. 16 — о «слогах *m*, *n*» в суахили (со ссылкой на исследование: A. N. Tucker, E. O. Ashton. Swahili Phonetics. «African Studies», vol. I, № 2—3, 1942) и стр. 11.

[32] См. «Народы Австралии и Океании», раздел «Меланезийские языки».

[33] Ср. в этой связи использование греками сочетания Кр для передачи начального k^w в малоазийских (в частности, карийских) собственных именах.

которого основывалась на фонологическом принципе — ср. также ниже, стр. 180).

Очевидно, объективное заключение о закономерностях строения и эволюции звуковых цепей может быть сделано лишь в том случае, если во внимание принимаются оба фактора — антропофонический и системный. Во всяком случае, отрицание понятия «труднопроизносимой группы согласных» (т. е. чисто «системный» подход к объяснению указанных звуковых явлений) наталкивается на серьезные возражения.

Структура звуковой цепи (следование элементов, относящихся к различным классам; следование элементов, относящихся к различным подклассам) может служить основанием для типологических сопоставлений — как в плане синхронии, так и в плане диахронии, причем оба плана оказываются теснейшим образом связанными между собой.

Структура звуковой цепи (например, в части сочетаемости согласных) в значительной степени обусловлена соотношением $N_C : N_V$; в этом смысле чрезвычайная структурная близость между финским и некоторыми африканскими языками, так поразившая в свое время Хакулинена[34], представляет собой весьма тривиальную типологическую закономерность. В той же связи следует рассматривать и тенденцию к упрощению консонантных групп по схеме $C_1C_2 > C\,(> C)$, чрезвычайно характерную для языков со сходным (в целом немногим превышающим единицу) отношением $N_C : N_V$. При исследовании языков, характеризующихся весьма разнообразными стечениями согласных, могут быть выделены наиболее устойчивые группы, обнаруживаемые даже в цепях тех языков, которые не допускают стечений согласных в пределах слога; такой чрезвычайно устойчивой группой может быть, например, сочетание $-RC(-)$. Естественно предположить, что подобные сочетания относительно древнее, чем сочетания иных типов, или, тем более, чем «нетипичные» сочетания, которые, как это было показано выше, вообще отличаются неустойчивостью.

Типологически сопоставимы не только отдельные языки, но и отдельные языковые стили (в понимании Щербы), причем звуковая цепь данного языка может оказаться близкой в структурном отношении звуковой цепи определенного стиля другого языка, тогда как оба языка могут иметь структурно различные звуковые цепи: в то же время звуковые цепи полного стиля данного языка и его разговорного стиля могут быть резко различны в структурном отношении.

Поскольку звуковые цепи со сходным соотношением $N_C : N_V$ имеют сходную структуру, сравнение разговорного стиля данного языка с полным стилем другого языка может пролить свет на ход эволюции первого языка в целом. Так, сравнение распространенного немецкого произношения и литературного английского произношения помогает лучше понять направление эволюции немецкого языка. Характерно в этом отношении сопоставление разговорного стиля французского языка и полного стиля русского языка: современное разговорное французское произношение характеризуется редукцией некоторых безударных гласных, что ведет к повышению консонантной насыщенности звуковых цепей. Для обоих языков характерны открытые слоги; выпадение предударных гласных во французском приводит к образованию начальных групп согласных, давно уже ставших характерными для русского языка, $N_C : N_V$ которого выше, чем во французском: *de la* ⩾ [dla], *devant* ⩾ [dvã], *de vue* ⩾ [dvü], *que la* ⩾ [kla], *cheveux* ⩾ [ʃvö], *semi* ⩾ [smi],

---

[34] Л. Хакулинен. Развитие и структура финского языка, I. М., 1953, стр. 17. Таблицы относительной частотности гласных и согласных, приводимые Хакулиненом (стр. 14), в ряде случаев не имеют под собой реальных оснований.

cependant ⩾ [spãdã] (образуются «типичные» группы со структурой *TR-SR-, ST-*).

В ликийской графике, видимо, зафиксирован процесс, соответствующий явлениям, наблюдаемым в беглой речи в современном русском языке, сопоставимый с некоторыми этапами развития славянских языков; важно отметить, что в ликийском языке подавляющее большинство конечных слогов открытые, а внутри слова интервокальные группы согласных соответствуют в большинстве случаев начальным группам, что, видимо, тоже свидетельствует о тенденции к построению открытых слогов; в ряде случаев группы согласных, включающих носовой слоговой звук, сопоставимы с соответствующим группам других языков, имеющих или имевших слоговые согласные.

Ликийские *pddedi; pzziti; clleima; kssbeze* сопоставимы в отношении способа образования с такими формами (приведенными в транскрипции Булича-Трахтерова, см. ВЯ, 1956, № 6, стр. 25), как [п̩дажди, п̩сажир, д̩лато, п̩сматри], причем структура этих образований может быть представлена и в виде *CəC̄V-*, *CəC̄CV-*, в особенности в отношении слов со смычными геминатами, что и давало грекам повод в ряде случаев прибегать к написаниям типа πάταρα, как, впрочем, и πταρά. Образования типа *cm̃ma, prñnawa, trm̃mila* сопоставимы с формами того же типа ([ф̩нари] и т. д.), причем ликийские *m̩* и *n̩* имеют специальные обозначения (как, между прочим, и примеры в приводимой здесь упрощенной транскрипции русских слов).

Естественно, подобного рода сопоставления могут быть использованы как для прогнозирования эволюции структуры данного языка (что, впрочем, едва ли может иметь какую-либо теоретическую или, тем более, практическую ценность), так и для реконструкции тех или иных особенностей мертвого или исторически незасвидетельствованного языка. Так, предположение о том, что индоевропейский язык некогда имел фонетическую структуру *CVCVCVCV...* при фонологической *CCCC...*, позволяет сопоставить его с разговорными стилями таких языков, как монгольский, калмыцкий или португальский, звуковые цепи которых резко отличны по структуре от соответствующих фонологических цепей; в обоих случаях эволюция, видимо, направлена на увеличение консонантной насыщенности звуковых цепей (в частности образование стойких моделей *TR*, *ST*, *RT* и т. д.). То состояние индоевропейского языка, которое характеризовалось становлением устойчивой слоговой структуры, сопоставимо с современным состоянием абазинского языка, в котором имеется всего одна или две гласные фонемы (что реконструируется также и для соответствующего состояния индоевропейского языка). Видимо, в индоевропейском, как и в абазинском, гласная фонема характеризовалась не столько противопоставлением другой гласной фонеме, сколько вообще отсутствию гласной фонемы; вместе с тем исключительную роль в качестве смыслоразличительного средства играло ударение, ср. абаз. *а́лхра* 'вычитать': *алы́хра* 'различать', ср. также лексические пары: абаз. *абара́* 'сохнуть' — *аба́ра* 'засуха', чрезвычайно напоминающее как в отношении формы, так и в отношении семантики известное противопоставление индоевропейских баритонных и окситонных образований[35]. В обоих языках основная смыслоразличительная нагрузка приходилась на согласные — отсюда слова типа *CVC*, образованные из слов *CV* + *CV*, а также разветвленная система смычных. Для более позднего этапа развития индоевропейского языка весьма показательно сопоставление с немецким: в обоих языках звуковые цепи характеризовались исключительно высокой консонантной насыщенностью,

[35] Сведения об абазинском языке почерпнуты главным образом из блестящей монографии А. Н. Генко «Абазинский язык» (М., 1955).

отсюда обилие закрытых слогов с чрезвычайно четкой структурой *-(S)(T)(R)-V-(R)(S)(T)* в немецком и *(S)(T)(R)-V-(R)(T)* в индоевропейском[36], отсутствие неприкрытых слогов — ср. нем. '- и и.-е. *H-* (ларингал) и т. д. Наличие в немецком слогов *CV*, представляющих единственно универсальную структуру слога в известных языках мира, приводит к необходимости принять схему слога *CV* для всех периодов развития индоевропейского языка[37].

Методика типологических сопоставлений, видимо, должна заключаться в следующем: текст в 500—1000 знаков, записанный фонетическим или близким к фонетическому письмом, подвергается всестороннему обследованию: выявляется соотношение $N_C : N_V$, подсчитывается количество *C* и *V* в начале и конце слов (если текст разделен на слова), выявляются структурные типы сочетаний элементов одного и того же класса, благодаря чему определяются подклассы, типичное отделяется от нетипичного и т. д. Подобный анализ текста может служить базой для достаточно четкой характеристики данного языка, отличающей его от всех других языков, даже близкородственных; с другой стороны, сравнение результатов, полученных на основе разных языков (или разных языковых стилей), может явиться основой той классификации языковых типов, о которой было сказано в начале настоящей работы.

ON THE STRUCTURE OF SOUND-CHAINS

Summary

The paper is concerned with the methodological problem of identifying vowels and consonants (class *V* and class *C*) in sound chains as well as with the problem of identifying elements belonging to subclasses (e. g. *S*, *R*, *T* of class *C*).

For a text in any given language recorded in phonetic transcription, or, for that matter, in letter writing, the identification of an element as belonging to the class of vowels or the class of consonants may be reduced to the identification of two large groups of three-member strings of succe) sive elements composed on the bases of the given text, namely, the group with the vowel in the center and the group with the consonant in the center.

The point is that even in a language in which three-member strings of successive sounds of the speech flow are formed after the optimal pattern *CCC*, *CCV*, *VCC*, *VCV*: *CVC*, *VVC*, *CVV*, *VVV*, certain types of soud combinations are predominant. One may identify elements of class *V* and class *C* by writing all possible triples of elements out of a given text (so as to place under a certain triple with the given central element all the other triples with the same central element) and by opposing the triples with the central-*V*- to the triples with the central-*C*-.

Another simple method is used to identify the subclasses.

These methods may be used for a study of an unknown text (i. e. for the purpose of deciphering): one may for instance decipher Karian, as well as any other unknown language recorded in letter-writing.

---

[36] В тексте слоги типа *(C)CCVCC(C)* встречаются крайне редко; слоги типа *CCVC* и *CVCC* встречаются много реже, чем слоги *CVC*; такое положение полностью соответствует тому, что было сказано в первой части настоящей работы относительно закономерностей взаимной дистрибуции элементов *C* и *V* в звуковой цепи.

[37] О различных периодах развития индоевропейского языка см. мои заметки «К истории индоевропейского генитива» (ВЯ, 1957, № 6) и «К проблемам реконструкции» («Филологические науки», 1958, № 3).

Е. Л. Гинзбург

НЕКОТОРЫЕ СЛУЧАИ НЕОДНОСОСТАВНОСТИ ПРОСОДЕМОИДОВ[1]

Опыт операционного анализа

I

АБСТРАКТНАЯ ПОСТАНОВКА ЗАДАЧИ

01. Задача исследования сводится к установлению характера отношений между заданными признаками, присущими каждому объекту наблюдаемой совокупности. Эта задача составляет часть определения эффективного пучка признаков, необходимого и достаточного для моделирования искомых лингвистических сущностей наблюдаемой совокупности. Иными словами, нужно определить существенные и важные связи признаков — такие связи, что их знание позволит дать характеристику любого объекта, принадлежащего, по определению, исследуемому множеству.

02. Современное языкознание исходит из признания в качестве своего предмета структуры плана выражения и структуры плана содержания, некоторые закономерности отношений между которыми определяются законом коммутации. Этот закон находит свое выражение в отношениях доминации[2] как между лингвистическими единицами, так и составляющими их признаками. Эти виды зависимостей характерны для отношений как внутри структур, так и между ними. Задача исследователя сводится к выявлению этих связей и установлению их характера.

03. Последние достижения языкознания основаны на признании того, что исследуемые лингвистические сущности выявляются лингвистом при восхождении от наблюдения формальных признаков или характеристик изучаемых лингвистических явлений к постулированию конструктов. Отражая объективно существующие противопоставления элементов структуры языка, эти признаки дают ключ к познанию лингвистических сущностей, к выявлению лингвистических инвариантов, к установлению имманентной структуры языка.

04. Современная фонология, изучающая структуру системы диакритических знаков, различает ингерентные и просодические дифференциальные признаки и исследует отношения между ними или их комбинациями. Грамматика, изучающая структуру системы знаков, должна

[1] Автор пользуется случаем, чтобы выразить глубокую признательность С. К. Шаумяну, В. А. Пурто, В. А. Виноградову за критические замечания и помощь в работе над рукописью.

[2] Об отношениях доминации см.: С. К. Шаумян. Трансформационная грамматика и теория классов слов. Тезисы докладов на конференции по структурной лингвистике, посвященной проблемам трансформационного метода. М., 1961, стр. 13 (допущения 3, 4).

исходить из грамматических дифференциальных признаков, комбинации которых образуют грамматические сущности или категории (ср., например, работы Р. Якобсона о падеже в славянских языках). Изучение отношений между р а з л и ч н ы м и структурами языка необходимо включает анализ отношений между признаками, определяющими их категории. Исследование, проводимое без рассмотрения этой связи, не является достаточно адекватным.

05. Отношениям между ингерентными признаками (и соответствующими категориями фонологии) и грамматическими признаками (и соответствующими категориями грамматики) посвящена богатая и разнообразная литература. Соответствующий раздел языковедения получил название морфонологии.

Значительно слабее разработан вопрос о связи п р о с о д и ч е с к и х и г р а м м а т и ч е с к и х д и ф ф е р е н ц и а л ь н ы х п р и з н а к о в и соответствующих им категорий.

КОНКРЕТНАЯ ПОСТАНОВКА ЗАДАЧИ

01. Задачу настоящей работы составляет рассмотрение отношений между признаками, определяющими, с одной стороны, п р о с о д е м о и д и п р о с о д е м у [3] и, с другой стороны, м о р ф е м н ы й с о с т а в с л о в а.

02. При некоторых допущениях признаками просодемоида можно признать одно- и неодносоставность, подвижность или неподвижность, префигированность, суффигированность, основность или флексивность (наличие просодемоида на приставке или суффиксе, основе или окончании) [4].

03. Понятия одно- или неодносоставности подразумевают выделение в определяемом просодемоидом слове одной или более одной вершины, одного или более одного слога. В данной работе как признаки, так и характеризуемые ими категории просодемы и просодемоида формально не определяются; некоторые аспекты содержания этой категории рассматриваются в разделе V, 3.

04. Наименее изучен вопрос об одно- и неодносоставности просодемоида [5]. Неодносоставность просодемоида в простых и производных словах (в большей части, кроме слов с *анти-*, *архи-*, *про-*, *супер-*, *транс-*, *ультра-* и т. п., характер которых подлежит выяснению) не наблюдается. В сложных и сложносокращенных словах (см. определения ниже) описываемый признак просодемоида широко распространен. Поэтому анализируемой совокупностью слов является множество сложных и сложносокращенных слов, определяемое соглашением о исследуемом материале.

---

[3] Понятие просодемы является экспликатом понятия ударения на уровне конструктов, а просодемоида — экспликатом ударения на уровне наблюдения. См. С. К. Ш а у м я н. Проблемы теоретической фонологии. М., 1962, стр. 64. Традиционное понятие ударения, таким образом, расщепляется на просодему и и просодемоид. Подобное осмысление термина «просодема» в корне отлично от ранее употреблявшихся значений этого термина. Ср., например, у Н. С. Трубецкого («Основы фонологии», М., 1960, стр. 222): «Под просодемой мы понимаем. . . слог в слогосчитающих языках и мору в моросчитающих языках». Здесь и далее термин «просодема» и «просодемоид» понимаются так, как эти понятия используются в фонологической концепции С. К. Шаумяна.

[4] Все перечисленные признаки всесторонне рассматриваются в обширной литературе по ударению в славянских языках. Исключение составляет признак одно- и неодноударности слова, экспликатом которого является одно- и неодносоставность просодемоида.

[5] Ср., например, у Р. И. Аванесова: «Вопрос о побочном, втором ударении слова совершенно не изучен» («Ударение в современном русском литературном языке», М., 1955, стр. 37).

Особый характер морфемного строя таких слов (наличие более чем одной корневой морфемы), способов образования, свидетельствующих об известном изоморфизме структур сложных слов и синтагм, внутренняя близость словосложения и инкорпорации, проблемы компрессии в языке и речи, энтропии и избыточности в структуре языка и многое другое определяет значение этой задачи.

10. Разработка алгоритмов автоматического морфологического анализа [6], проблем морфемного анализа [7] и основ теоретической фонологии [8] не только делают эту задачу своевременной, но и позволяют в основном разрешить ее (хотя и не полностью — из-за отсутствия данных экспериментальной фонетики).

20. Вслед за соглашениями о характере, типе и аспектах исследования приводится определение сложного слова, затем рассматривается логическая структура переменных признаков, возможность автоматизации этой лингвистической задачи, характер процедур выписывания алгоритма, отношение между классами и алгоритмом их выделения, некоторые требования к алгоритму. В разделе IV приводится алгоритм соотнесения переменных и описываемого признаков, произведен анализ полученных классов. Определяются понятия составного производного признака корреляционной системы, систем ненаправленной и направленной корреляции, систем независимых отношений.

В заключение данные анализа рассматриваются в свете двухступенчатой теории структуры языка С. К. Шаумяна.

II

1. СОГЛАШЕНИЕ О ТИПЕ ИССЛЕДОВАНИЯ

Настоящая работа представляет собой системную генерализацию [9] эмпирических данных, определенных соглашением о материале исследования, и являет собой один из этапов выдвижения гипотезы, на основе которой дедуктивным путем может быть построена модель — лингвистическое исчисление сложных слов.

Системной генерализацией назовем описание некоторого множества языковых единиц, базирующееся на соотнесении значений переменных и описываемого признаков, принадлежащих каждому объекту наблюдаемой совокупности объектов.

Системной генерализации предшествует простая генерализация, которая есть установление значений переменных признаков и описываемого признака, следующих из формального определения наблюдаемого объекта; приписывание каждому объекту пучка значений признаков; классификация подобных пучков значений.

---

[6] И. А. Мельчук. Морфологический анализ при машинном переводе (преимущественно на материале русского языка). «Проблемы кибернетики», вып. 6, 1961, стр. 207—276. В библиографии указаны все наиболее важные работы по автоматическому морфологическому анализу текста, изданные до 1961 г. Более новые работы зарубежных авторов по морфологическому анализу при машинном переводе приведены в кн.: Р. Д. Равич. Библиография зарубежных работ по машинному переводу (1960—1961). Предварительные публикации Института языкознания АН СССР. М., 1962.

[7] И. И. Ревзин. О логической форме лингвистических определений. «Применение логики в науке и технике», М., 1961, стр. 140—148.

[8] См. работы: С. К. Шаумян. Проблемы теоретической фонологии; его же. Структурная лингвистика как имманентная теория языка. М., 1958; его же. Двухступенчатая теория фонемы и дифференциальных элементов. ВЯ, 1960, № 5.

[9] См. С. К. Шаумян. Генерализация и постулирование конструктов. Тезисы совещания по математической лингвистике. Л., 1959; его же. Преобразование информации в процессе познания и двухступенчатая теория. Тезисы докладов на конференции по обработке информации, машинному переводу и автоматическому чтению текста. М., 1961.

Системная генерализация является высшим типом генерализации, так как рассматриваемая совокупность характеризуется в отношении одного из признаков, который в свою очередь анализируется на основе соотношений с другими признаками в различных их последовательностях.

2. СОГЛАШЕНИЕ О ХАРАКТЕРЕ ИССЛЕДОВАНИЯ

01. Исследование носит операционный характер [10]. Искомые лингвистические сущности определяются на основе задаваемой процедуры исследования. Последняя состоит в установлении наличия или отсутствия заданных формальных признаков, приписывании объектам значений этих признаков и нахождении полученных комбинаций значений признаков, таких комбинаций значений признаков, которые прогнозировали бы значения описываемого признака.

02. В силу того, что в основе описания [11] — наиболее простые и универсальные категории, а все остальные признаки получены логическим путем, операционное исследование обеспечивает наиболее целостное изучение элементов системы, наиболее полное отражение сущности изучаемого явления, многообразия связей отдельного и общего, взаимосвязи элементов анализируемой совокупности. В силу логической прозрачности алгоритма обеспечивается, кроме того, наиболее эффективная и экономная компенсация ошибок.

030. Для операционного исследования существенно понятие оператора.

031. Установление соответствия между элементами двух множеств X и Y (где X — множество признаков x, установленных логическим или интуитивным путем и соответствующих объекту по определению, а Y — множество признаков y, принадлежащих каждому объекту исследуемой совокупности и установленных в результате наблюдения над ним), такого, что каждому x из X относится некоторый y из Y, назовем операцией. Если эта операция составляет содержание некоторого правила какого-нибудь алгоритма, то назовем это правило оператором [12].

032. Из определения оператора следует необходимость построения системы признаков, соответствующих объекту анализа по его определению. Этой задаче и посвящена большая часть раздела III.

3. СОГЛАШЕНИЕ ОБ АСПЕКТАХ ИССЛЕДОВАНИЯ

01. В настоящей работе структура сложных слов исследуется в статическом аспекте [13], т. е. устанавливаются характеристики, определяющие для некоторого установившегося состояния описываемой совокупности слов зависимость между переменными признаками, принадлежащими всякому объекту этой совокупности. Здесь переменный признак понимается так, как это принято в физике. В данном случае это «односоставность — неодносоставность просодического элемента» и признаки, характеризующие морфемный состав сложного слова. Цель такого исследования — определить параметры, характеризующие описываемую

---

[10] См. С. К. Шаумян. Операционные определения и их применение в фонологии. «Применение логики в науке и технике», М., 1961, стр. 149.

[11] О понятии «описание» см. И. А. Мельчук. О стандартной форме и количественных характеристиках лингвистических описаний. В. Я., 1963, № 1.

[12] См. И. А. Мельчук. О стандартных операторах для алгоритма автоматического анализа русского научного текста. «Машинный перевод. Труды Института ТМ и ВТ АН СССР», М., 1961. См. также: О. С. Кулагина. Операторное описание алгоритмов перевода. «Машинный перевод и прикладная лингвистика», № 2 (9), № 3 (10).

[13] Излагаемое сводимо к различению системы и процесса у Л. Ельмслева, см. «Пролегомены к теории языка» («Новое в лингвистике», М., 1960, стр. 281 и далее). Это же по существу и у Н. Арутюновой («Очерки по словообразованию в современном испанском языке», М., 1961, стр. 10—15, 48—49).

систему; их взаимоизменение и взаимообращение — предмет динамического анализа. Частным случаем последнего является генетический анализ, в результате которого устанавливается логическое следование по степени производности или основности. Динамический анализ здесь не проводится, он составляет особую задачу.

02. Сложные слова можно изучать с точки зрения отношения к соответствующим им словосочетаниям (т. е. в плане порождения) или в сопоставлении с простыми и производными словами (т. е. в плане их морфемного состава). Сложные слова, таким образом, становятся объектом статического и динамического изучения, которые дополняют и уточняют друг друга.

03. При генерализации различение статического и динамического в сложном слове находит свое выражение в двух различных процедурах приписывания признаков, определения, группировки, классификации и систематизации описываемых объектов.

04. Введение требования различать статическое и динамическое опирается не только на логические и эпистемологические соображения, но и на практику некоторых крупных лингвистов [14].

4. СОГЛАШЕНИЕ О ХАРАКТЕРЕ АНАЛИЗИРУЕМОЙ СОВОКУПНОСТИ СЛОЖНЫХ СЛОВ

01. Анализ проводится на материале русского языка. В качестве информационного материала используются сложные и сложносокращенные слова, отвечающие приводимым ниже определениям и зарегистрированные в словаре-справочнике «Русское литературное произношение и ударение» под редакцией Р. И. Аванесова и С. М. Ожегова. Выборка производилась по изданию 1960 г.

02. Произведен анализ 1293 слов. Из них 818 — неодноударные (буквы от П до Я), 475 — одноударные (буквы Р, Т, У). В рассмотренной совокупности 498 слов — имена существительные, 769 — имена прилагательные, 10 — наречия, 16 — глаголы.

03. Допущение. Предполагается, что на основе одного из алгоритмов морфологического анализа, например алгоритма И. А. Мельчука, каждому анализируемому слову приписана в некоторой системе символов информация о морфемной структуре и, кроме того, о просодических элементах этого слова. Эта информация представлена некоторой заранее определенной последовательностью значений признаков, каждый из которых может принимать значение + или —. В полученных таким способом определениях приводятся значения всех признаков, характеризующих данное множество объектов. Поскольку значения признаков могут быть установлены лишь в результате осуществления некоторой совокупности операций над исследуемым множеством объектов, такие определения носят операционный характер.

04. Ограничения. В работе принимались тождественными случаи грамматического и словообразовательного варьирования, такие, как, например, варьирование существительных по роду или глаголов по залогу.

Из рассмотрения при анализе исключаются также следующие группы слов:

1. Слова, в начале которых морфемы *анти-*, *архи-*, *контр-*, *поли-*, *после-*, *пост-*, *псевдо-*, *теле-*, *транс-*, *ультра-*, *уни-*, *экс-*, *экстра-* и т. д.: *полисинтетический*, *послеуборочный*, *псевдонаучный*, *телевещание*, *трансарктический*, *ультракороткий*, *униформа*, *экс-министр*, *экстра-почта*

[14] Ср., например, у Л. А. Булаховского двухсторонний анализ сложных слов, который, к сожалению, проводится недостаточно последовательно (Л. А. Булаховский. Курс русского литературного языка, т. I. Киев, 1952, стр. 192—194).

и т. д. Характер этих морфем неопределенен и заслуживает особого рассмотрения. Поскольку бóльшая часть из них ударна наряду с той морфемой, к которой они присоединяются, рассмотрение их было бы целесообразным после анализа сложных слов, отвечающих определению сложного слова.

2. Звукоподражания типа *тик-так*, *трах-тарарах*, *трень-брень*, *тук-тук*, *хрю-хрю*, *цып-цып* и т. д. Их особый характер не раз отмечался в литературе. Анализ дубликатов вида *идти-идти*, *тьма-тьмущая*, *рад-радехонек* и т. д. позволит, думается, подойти к более углубленному анализу этих образований.

3. Слова с гаплологическим упрощением основ, например, *трагикомедия*, *морфонема* и т. д.

4. Случаи типа *сеньорен-конвент*, *ринодерма*, *рейнвейн*, *порттабак*, *папье-маше*, *тау-сагыз*, *терц-мажор*, *форс-мажор*, *цирлих-манирлих*, *шевинг-процесс*, ассимилированность которых близка к нулю; но само это утверждение может быть установлено лишь после анализа всей системы сложных слов.

5. Слова, начинающиеся с морфем *радио-*, *авиа-* и т. п. Характер этих морфем в статической системе сложных слов является неопределенным. Думается, что признаки, установленные на большой совокупности сложных слов, не вызывающих сомнения относительно характера их морфемного строя, позволят дать обоснованное решение этой проблемы.

6. Слова на *сам-*, например, *сам-восемь*, *сам-двадцать*, *сам-девять*, *сам-друг* и т. п.

III

1. ОПРЕДЕЛЕНИЕ СЛОЖНОГО СЛОВА

010. При определении сложного слова опираемся на рекурсивное определение понятия «морфема», данное И. И. Ревзиным[15]. Приведем его полностью, так как для настоящей работы оно является базовым.

011. «Алгоритм $\mathfrak{A}$, построенный нами для выделения морфемы, распадается на несколько этапов:

I этап. Все словоформы языка (или той части языка, которая исследуется нами) разбиваются на семейства, т. е. на непересекающиеся классы элементов, эквивалентных между собой в том смысле, что замена одного на другой переводит любую осмысленную фразу в осмысленную, а любую бессмысленную фразу — в бессмысленную. Слова в каждом семействе располагаются одно под другим в виде списка.

II этап. Любое деление слова на две части — правую и левую — называется разделом. Получить разделы данного слова означает последовательно представить слово w в виде $s_i + d_{n-i}$, где индекс обозначает число букв в s (левой части) или d (в правой части), n — число букв в слове, $i = 0, 1, 2, \ldots, n$. Разделы, для которых $i = 0$ и $i = 1$, считаются крайними.

III этап. Если при замене части p_1 на p_2 ($p = s_i$ или $p = d_{n-1}$) каждое слово, в котором присутствовало p_1, остается в своем семействе, то мы будем говорить, что p_2 подчиняет p_1 ($p_2 \supseteq p_1$).

Ряд элементов $p_1, p_2, \ldots, p_n$, таких, что $p_1 \subseteq p_2 \subseteq \ldots \subseteq p_n$, составляет парадигматический ряд. Если $p_1 \subseteq p_2$ и $p_1 \supseteq p_2$, то $p_1 \sim p_2$.

IV этап. Раздел, который не является крайним для всех элементов ряда и при котором для какой-то части данного слова w образуется парадигматический ряд, насчитывающий наибольшее количество элементов, называется главным разделом данного слова. (Если образуется несколько парадигматических рядов с наибольшим числом элементов, то главным считается самый левый раздел.)

[15] И. И. Ревзин. О логической форме лингвистических определений, стр. 144—147.

Парадигматический ряд элемента p_1 при главном разделе называется основным классом элемента p_1 данного слова w. Каждый элемент основного класса считается основой.

V этап. Два основных класса считаются пересекающимися, если они имеют хотя бы один общий элемент. Общая часть двух пересекающихся основных классов, если она не пуста, называется регулярной частью данного основного класса по отношению к другому основному классу. Часть основного класса называется просто регулярной, если она является наибольшей из частей, регулярных по отношению к любому основному классу, пересекающемуся с данным. Основа, относящаяся к регулярной части, называется регулярной.

VI этап. Слово, при исключении которого увеличивается число эквивалентных между собой элементов в данном основном ряду, называется словом с нерегулярным членением. Такие слова из рассмотрения исключаются.

VII этап. Возьмем все слова с данной регулярной основой; вычленив основу, получим в каждом из них остаток, называемый формантом. Формант называется префиксом, если он стоит слева от основы, и суффиксом, если стоит справа от основы.

Все слова вида $p_i + f$, где p_i подчиняет данную регулярную основу, а f — формант данной регулярной основы, анализируются так же.

VIII этап. Морфемой считается любой формант и любая основа, в которых не укладывается целиком другая морфема».

0121. Из сказанного всякое слово можно представить в следующей форме: $pr + T + sf$, где pr — префикс, sf — суффикс, T — основа.

0122. Суффикс, в который целиком укладывается иной суффикс, не тождественный данному, назовем составным. Он мог бы быть представлен в следующей форме: $sf_\alpha \supset sf_\beta$, где sf_β — включаемый суффикс, а sf_α — включающий суффикс, $\supset$ — знак включения. Часть составного суффикса, остающуюся при исключении части, тождественной вкладывающемуся суффиксу, назовем элементарным суффиксом тогда и только тогда, когда нет ни одного другого суффикса, который бы вкладывался в него, кроме ему тождественного.

0123. Элементарный суффикс, справа от вкладывающейся части, предшествующей ему, назовем флексией (ф).

0124. Всякий суффикс слева от флексии, составной или элементарной, назовем истинным суффиксом или просто суффиксом.

Истинный суффикс представляет собой множество элементарных суффиксов, мощностью от 0 до n, где $n \ll \infty$. Аналогично флексия может рассматриваться как множество, мощность которого от 0 до n, где $n < \infty$.

Аналогично истинному суффиксу префикс тоже может быть элементарным и составным.

0125. Исходя из сказанного, формула всякого слова могла бы приобрести следующий вид:

$$[Pr\,(pr_\alpha \supset pr_\beta \supset \ldots \supset pr_\omega) + T + Sf\,(sf_\alpha \supset sf_\beta \supset \ldots \supset sf_\omega)]\,(\text{ф}). \qquad (2)$$

0126. Часть слова, остающаяся после вычленения из него флексии, называется основной морфемой (ОМ). Основная морфема может быть представлена в следующей форме:

$$[Pr\,(pr_\alpha \supset pr_\beta \supset \ldots \supset pr_\omega) + T + Sf\,(sf_\alpha \supset sf_\beta \supset \ldots \supset sf_\omega)]\,(\text{ф}). \qquad (3)$$

013. Определение. Сложным словом назовем такое слово, в котором при выделении одной основной морфемы останется по крайней мере еще одна основная морфема.

02. Пример. Доказать, что *водянисто-голубоватый* — сложное слово.

Пусть задан следующий кортеж слов: *водный*, *водянистый*, *водянисто-голубоватый*, *водяный*, *волосяный*, *волосянистый*, *волосистый*, *железистый*, *искристый*, *овсяный*, *рамчатый*, *росяный*, *рыжеватый*, *синеватый*, *струйчатый*.

Выпишем все слова, начинающиеся сочетаниями букв *в-о-д-* и *в-о-л-о-с-*, кроме рассматриваемого:

C_1	C_2
водный	*волосяный*
водянистый	*волосянистый*
	волосистый

Знаком $\sqrt{\ }$ обозначим раздел.

В семействе C_1, C_2 главный раздел пройдет следующим образом:

1) *водн*$\sqrt{\ }$*ый*
2) *водянист*$\sqrt{\ }$*ый*
3) *волосян*$\sqrt{\ }$*ый*
4) *волосянист*$\sqrt{\ }$*ый*
5) *волосист*$\sqrt{\ }$*ый*

Обозначенный раздел является главным, так как перемена места любых элементов левой части не выводит из этого семейства ни одного слова, при этом парадигматический ряд в данном случае наибольший.

Исключив 1-й, 3-й и 5-й элементы семейства C_1, C_2, оказывается возможным провести главный раздел левее, после частей *вод-* и *волос-*. По определению морфемы, более левый раздел оказывается главным. Отсюда

вод$\sqrt{\ }$*янистый*
волос$\sqrt{\ }$*янистый*,

что подтверждается принадлежностью к этому семейству также

дерев$\sqrt{\ }$*янистый*
лед$\sqrt{\ }$*янистый*
трав$\sqrt{\ }$*янистый*.

С учетом первого членения получаем

вод$\sqrt{\ }$*янист*$\sqrt{\ }$*ый*
волос$\sqrt{\ }$*янист*$\sqrt{\ }$*ый*
дерев$\sqrt{\ }$*янист*$\sqrt{\ }$*ый*
лед$\sqrt{\ }$*янист*$\sqrt{\ }$*ый*
трав$\sqrt{\ }$*янист*$\sqrt{\ }$*ый*.

Итак, в результате проведенных операций можно выделить основы *вод-*, *волос-* и составной суффикс *-янистый*, который разлагается на флексию *-ый* и истинный суффикс *-янист-*.

Докажем, что суффикс *-янист-* — истинный составной суффикс.

Аналогичный анализ совокупности слов *водяный*, *волосяный*, *росяный*, *овсяный* и им подобных или совокупности слов *волосистый*, *железистый*, *искристый* верифицирует наше утверждение как истинное, при этом для подобного утверждения вполне достаточно провести анализ одной из совокупностей. Последующее сравнение членений приводит к разложению суффикса *-янист-* на два элементарных суффикса *-ян-* и *-ист-*. Оба они являются истинными.

Тогда, по определению, *водянист-* является основной морфемой.

Рассмотрение семейства

водянистый
водянисто-голубоватый

приводит к выделению в последнем этой же основной морфемы. Если бы в этом слове оказалась еще одна основная морфема, то его можно было бы признать сложным.

Аналогичная операция, проведенная с семействами

струйчатый	*рыжеватый*
рамчатый	*голубоватый*
тесемчатый	*синеватый*

приводит к выделению основной морфемы *голубоват-*, в состав которой входит составной истинный суффикс *-ов-ат-*.

Рассмотрение семейства *голубоватый* и *водянисто-голубоватый* приводит к выделению второй основной морфемы. Тогда, по определению, *водянисто-голубоватый* является сложным словом.

03. Выделение в составе слова основ и формантов, основных морфем и флексий, суффиксов и префиксов, а в составе последних — элементарных суффиксов, установление отношений подчиненности между морфемами назовем морфемным анализом.

04. Состав слов, устанавливаемый морфемным анализом, назовем морфемным составом слова.

05. Определим некоторые понятия, относящиеся к морфемному составу сложного слова.

051. Часть сложного слова, остающуюся после вычленения из него флексии, назовем сложной основной морфемой этого слова. В слове *водянисто-голубоватый* ею является *водянисто-голубоват-*.

052. Если для данной сложной основной морфемы есть такая сложная основная морфема, что, не будучи тождественна первой, содержит те же самые простые основные морфемы, то часть, остающаяся при вычитании из первой сложной основной морфемы второй, называется формантом сложного слова, к которому принадлежит первая сложная основная морфема.

Форманты сложного слова могут аналогично формантам простого слова быть расчленены на префиксы, инфиксы и суффиксы. Так как определение этих категорий является побочным по отношению к данной работе, то мы его опускаем.

Суть приведенных определений состоит в том, что они исходят из наличия некоторого алгоритма, который можно применить к словам языка L с целью выделения определяемых категорий.

053. Часть сложного слова между двумя основными морфемами, не принадлежащая ни одной из них, называется соединительной морфемой, представляющей множество элементов от 0 до 1 и репрезентируемой одним из алломорфов $\varnothing$,*-о*, *-е* и др., где $\varnothing$ — нулевая морфема.

054. Допущение. Соединительная морфема принадлежит сложной основной морфеме. Если сложная основная морфема содержит $r+1$ простую основную морфему, то r-ая соединительная морфема относится к r-ой простой основной морфеме.

06. Введем следующие обозначения:

$K=\{A, B, \ldots, X\}$ — основы;

$k=\{k_i, k_j, \ldots, k_r\}$ — соединительные морфемы;

$\Phi=\{\phi_i, \phi_j, \ldots, \phi_r\}$ — флексии;

$l=\{a_i, b_i, c_j, \ldots, x_j\}$ — префиксы, если малая латинская буква перед большой латинской буквой;

$n=\{a_i, b_j, c_j, \ldots, x_j\}$ — суффиксы, если малая латинская буква после большой латинской буквы;

$\mathfrak{C}=\{\mathfrak{C}', \mathfrak{C}'', \ldots, \mathfrak{C}^{r+1}\}$ — основные морфемы;

| |, [ ], ( ) — границы морфем.

На основе введенных обозначений попытаемся представить сложное слово в виде канонической формулы.

07. По определению сложного слова, оно может быть представлено по крайней мере следующей формулой:

$$(\mathfrak{C}'k_j\mathfrak{C}'')\,|\phi_i| \tag{1}$$

Проведем подстановку $S^{aBc}_{\mathfrak{C}'}S^{dEf}_{\mathfrak{C}''}$ (1). Получим

$$\vdash(aBck_jdEf)\,|\phi_i| \tag{2}$$

Произведем подстановку $S^{lMnk_jaBc}_{aBc}$ (2), где lMn — последовательность любых букв из алфавитов l, K, n соответственно. Применив эту подстановку r раз, получим развернутую формулу сложного слова

$$\vdash(aBck_1gHik_2\ldots k_rdEf)\,|\phi_i| \tag{3}$$

Такое представление сложного слова легко привести к более удобной форме

$$\vdash\{[(lMn)(k)]^r(dEf)\}\,|\phi_i|; \tag{4}$$

r может быть интерпретирована и как величина, которая меньше мощности множества $\mathfrak{C}=\{\mathfrak{C}', \mathfrak{C}'', \ldots, \mathfrak{C}^{r+1}\}$ для данного слова на единицу [16].

При $r=0$ формула сложного слова обращается в формулу простого слова и приобретает следующий вид:

$$\vdash(dEf)\,|\phi_i|. \tag{5}$$

080. Ниже делается попытка дать истолкование понятия «сложносокращенное слово». Следует подчеркнуть, что это истолкование никоим образом не является строгим определением. Скорее это один из возможных подходов к выработке требуемого определения.

081. Соглашение. Назовем усеченной основой такую основу, для которой

1) есть основа, подчиняющаяся ей, т. е. не тождественная ей, но включающая ее как часть;

2) разность этих основ не тождественна ни одной другой морфеме или, точнее, не принадлежит ни к одному из классов морфем;

3) в остатке слова, из которого вычленена описываемая морфема, есть по крайней мере один элемент, который не принадлежит к формантным классам морфем.

082. Определение. Сложное слово, в котором есть хотя бы одна усеченная основа, назовем сложносокращенным словом.

083. Если усеченная основа является последней, сложносокращенное слово назовем сложносокращенным с конечным усечением.

Если усеченная основа является начальной, то сложносокращенное слово назовем сложносокращенным с начальным усечением.

084. Как и другие сложные слова, сложносокращенные слова характеризуются такими признаками, как, например, наличие соединительных морфем, величина r и т. д.

---

[16] Характеристика пределов (прежде всего верхнего) изменения величины r представляется крайне важной в свете гипотезы Ингве. См. V. H. Yngve. A model and a hypothesis for language structure. «Research Laboratory of Electronics and department of modern languages, Massachusetts Institute of Technology», technical report 369. Cambridge, Massachusetts, 1960, p. 444—466. И. И. Ревзин. Интересная гипотеза Ингве. «Машинный перевод и прикладная лингвистика», 1961, № 5 (73—80); И. И. Ревзин. Модели языка. М., 1962, стр. 47, 52.

2. ПЕРЕМЕННЫЕ ПРИЗНАКИ

00. Соглашение. Понятие «признак» не определяется и является исходным.

01. Определение. Признаки, подлежащие исследованию и принадлежащие всякому элементу анализируемой совокупности, называются описываемыми (*Fd*).

02. Определение. Признаки, принадлежащие всякому объекту анализируемой совокупности слов и используемые для исследования описываемых признаков, называются переменными (*Fp*)[17].

03. Описываемый и переменные признаки признаются заданными. Задача сводится к установлению отношений между ними.

04. Описываемым признаком в работе является одно- или неодносоставность просодемоида сложного слова. Переменными являются признаки, характеризующие морфемный состав сложного слова.

041. Исходя из определения сложного слова и соглашения о типе, характере и аспектах исследования, строится следующая система признаков:

0411. $r=1$ или $r>1$. В другой формулировке, мощность множества $\mathfrak{C}=\{\mathfrak{C}', \mathfrak{C}'', \ldots, \mathfrak{C}^{r+1}\}$ равна двум или больше двух. Интерпретация: «В сложном слове две основные морфемы?» Ср. *паровоз, электрополотер, славяно-греко-латинский, вектор-электрокардиограф.*

0412. $k=0$ или $k=1$. Интерпретация: «В сложном слове (есть) нет соединительной морфемы». Ср. *пила-рыба, машбюро,* но *пилорама.*

0413. $aBc \subseteq dEf$ или $\overline{aBc \subseteq dEf}$ ($dFf \subseteq aBc$ или $\overline{dEf \subseteq aBc}$). Интерпретация: «Включает ли одна из основных морфем сложного слова другую основную морфему этого же слова?». Частный случай: $aBc=dEf$. Интерпретация: «Тождественны ли основные морфемы?» Ср. *тьма-тьмущая, идти-идти,* но *жаронепроницаемый.*

0414. Примечания.

1. Из системы признаков указываются те и только те признаки, которые оказались необходимыми и достаточными для решения задачи. Например, опускается следующий признак: в $(aBckdEf)|\phi_i|\ c=\phi_j \in \Phi_i$, где $\Phi_i=\{\phi\}$. Интерпретация: «В состав непервой основной морфемы в сложном слове входит ли такая часть, которая омонимична одной из флексий для этой основы при ее независимом использовании?» Ср. *царь-птица* и *царь, пила-рыба* и *пила* и др.

2. Описание признаков в 0412 и 0413 подразумевает $r=1$. Сформулированные признаки действительны и для случая $r \geqslant 2$. Изменилась бы лишь формулировка. Так, описание в 0412 могло бы быть переформулировано следующим образом: для сложного слова с мощностью n множества $\mathfrak{C}=\{\mathfrak{C}', \mathfrak{C}'', \ldots, \mathfrak{C}^n\}$, где $n \ll \infty$,

$$k_1=0 \text{ или } k_1=1,$$
$$k_2=0 \text{ или } k_2=1,$$
$$\ldots\ldots\ldots\ldots$$
$$k_{n-1}=0 \text{ или } k_{n-1}=1.$$

Введение подобных признаков усложнило бы принципиальную схему предлагаемой процедуры, но не расширило бы области ее применения, так как по отношению к признаку «односоставность — неодносоставность просодемоида» при $r \geqslant 2$ большинство переменных признаков иррелевантно.

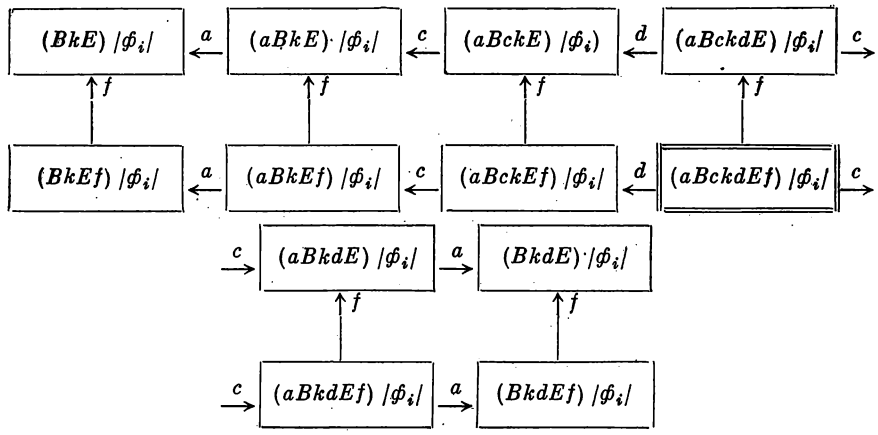
[17] Требования, которым должны отвечать такие признаки, см. в работе: P. Garvin. From model to procedure. Proceedings of the national symposium on machine translation. New York, 1961, стр. 368. «To be useful the properties set forth in a definitional model must be general enough not to prejudge the results, and specific enough not to be trival. They must allow the differentiation of one's object of cognition from other similar or related objects, and they must provide suitable points of departure for a procedure. They must allow the deduction of the equivalence principle and relevance criterion proper to one's own field of study».

05. В зависимости от того, какие значения получают элементы a, c, d и f в сложном слове $[(aBckdEf)\,|\,\phi_i\,|]$ и в простых словах $[(aBc)\,|\,\phi_j\,|]$ и $[(dEf)\,|\,\phi_k\,|]$, мог бы быть представлен следующий набор цепочек. Знак $\vdash$ следует читать: «В языке L есть такое слово, которое является материальным субстратом данного символа».

1. $\vdash (aBckdEf)\,/\phi_i/$	2. $\vdash (aBc)\,/\phi_i/$	3. $\vdash (dEf)\,(\phi_i)$
4. $\vdash (aBckdE)\,/\phi_i/$	5. $\vdash (aB)\,/\phi_i/$	6. $\vdash (dE)\,/\phi_i/$
7. $\vdash (aBckEf)\,/\phi_i/$	8. $\vdash (Bc)\,/\phi_i/$	9. $\vdash (Ef)\,/\phi_i/$
10. $\vdash (aBckE)\,/\phi_i/$	11. $\vdash (B)\,/\phi_i/$	12. $\vdash (E)\,/\phi_i/$
13. $\vdash (aBkdEf)\,/\phi_i/$	14. $\vdash (BkdEf)\,/\phi_i/$	15. $\vdash (aBkEf)\,/\phi_i/$
16. $\vdash (aBkE)\,/\phi_i/$	17. $\vdash (BkEf)\,/\phi_i/$	18. $\vdash (BkE)\,/\phi_i/$

Процесс получения производных цепочек для сложного слова мог бы быть представлен схемой (табл. 1), где $\xrightarrow{l}$ — символ оператора стирания в исходной цепочке символа l.

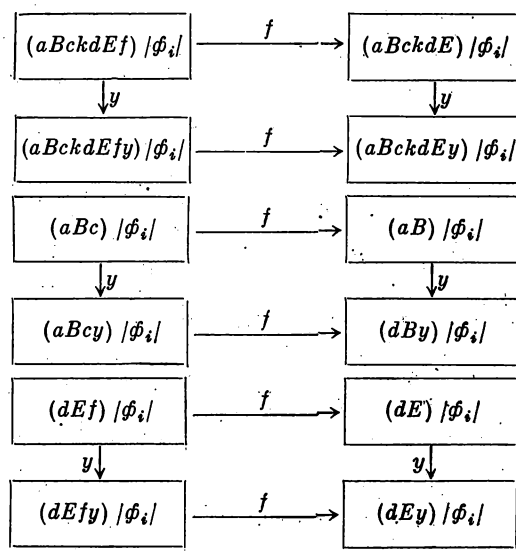
Таблица 1



В двойную рамку заключено исходное слово. Производство цепочек от $(aBc)\,|\,\phi_i\,|$ и $(dEf)\,|\,\phi_i\,|$ могло бы быть представлено в такой же форме и экономии места ради не приводится.

По соображениям, изложенным в примечании 1 (0414), в качестве переменных признаков используются соответствия сложных слов первым шести исходным и производным цепочкам символов и производным от них в результате применения операций, описанных в 052.

Таблица 2



052. Применим к первым шести цепочкам символов из списка в 05 операцию замещения позиции между $|\,\phi_i\,|$ и основной морфемой некоторой буквой y, такой, что ее прибавление к символам сложных основных морфем дает также символы сложных основных морфем, не тождественных первым. Иными словами, $y \neq 0$. Так как символ y расположен справа от большой латинской буквы, то интерпретируем его

как новый элементарный суффикс, составляющий с прежними, не равными нулю, новый составной суффикс. Схема (табл. 2) изображает порождение новых цепочек. В приводимой схеме $\xrightarrow{f}$ — символ операции по стиранию символа f из цепочки, откуда выходит стрелка и которую назовем исходной. $\downarrow y$ — символ операции прибавления буквы y к слову, в которое стрелка входит. Буква y замещает пустую позицию между основной морфемой и флексией в слове, откуда выходит стрелка.

06. Из сказанного выше следует, что в дополнение к первым трем признакам (0411—0413) могут быть сформулированы еще 12 признаков, основанных на следующих противопоставлениях:

1.		$(aBckdEf)\xrightarrow{}/\phi_i/$	$\vdash (aBckdEf)/\phi_i/$
2.		$(aBckdEf)/\phi_i/$	$\vdash (aBckdE)/\phi_i/$
3.		$(aBckdEf)/\phi_i/$	$\vdash (aBckdEfy)/\phi_i/$
4.		$(aBckdEf)/\phi_i/$	$\vdash (aBckdEy)/\phi_i/$
5.	$(dEf) \subset$	$(aBckdEf)/\phi_i/$	$\vdash (dEf)/\phi_j/$
6.	$(dEf) \subset$	$(aBckdEf)/\phi_i/$	$\vdash (dE)/\phi_i/$
7.	$(dEf) \subset$	$(aBckdEf)/\phi_i/$	$\vdash (dEy)/\phi_i/$
8.	$(dEf) \subset$	$(aBckdEf)/\phi_i/$	$\vdash (dEfy)/\phi_i/$
9.	$(aBc) \subset$	$(aBckdEf)/\phi_i/$	$\vdash (aBc)/\phi_j/$
10.	$(aBc) \subset$	$(aBckdEf)/\phi_i/$	$\vdash (aB)/\phi_i/$
11.	$(aBc) \subset$	$(aBckdEf)/\phi_i/$	$\vdash (aBy)/\phi_i/$
12.	$(aBc) \subset$	$(aBckdEf)/\phi_i/$	$\vdash (aBcy)/\phi_i/$

0700. На основе противопоставлений 1—12 могли бы быть сформулированы следующие 12 переменных признаков.

0701. Для $[(aBckdEf)|\phi_i|]\vdash[(aBckdEf)|\phi_i|]$ такое, что при тождестве основных сложных морфем ϕ_i и ϕ_j не только не тождественны, но таковы, что окрестности слов, к которым они относятся, различны, т. е. $\phi_i \neq \phi_j$, так как $\phi_i \in \Phi_l$, а $\phi_j \in \Phi_m$ при $l \neq m$. Интерпретация: «В языке L для сложного слова $\mathfrak{A}$ (левая часть противопоставлений) существует такое сложное слово $\mathfrak{B}$ (правая часть противопоставлений), что окрестности слов $\mathfrak{A}$ и $\mathfrak{B}$ не тождественны». Ср. *Железноводск — железноводский, целеустремленный — целеустремленно.*

0702. Для $[(aBckdEf)|\phi_j|]\vdash[(aBckdE)|\phi_i|]$ независимо от того, эквивалентны ли окрестности слов $\mathfrak{A}$ и $\mathfrak{B}$ при $f \neq 0$. Интерпретация: «В языке L для сложного слова $\mathfrak{A}$ существует такое сложное слово $\mathfrak{B}$, что $\mathfrak{A}$ включает $\mathfrak{B}$ как составную часть». Ср. *вагоностроительный — вагоностроитель, кинематографический — кинематограф.*

0703. Для $[(aBckdEf)|\phi_i|]\vdash[(aBckdEfy)|\phi_j|]$ при $y \neq 0$. Интерпретация: «В языке L для сложного слова $\mathfrak{A}$ существует такое сложное слово $\mathfrak{B}$, что ОМ$\in\mathfrak{B}$ включает в качестве составной части ОМ$\in\mathfrak{A}$, при этом $y \neq 0$, т. е. ОМ этих слов не тождественны». (Здесь символ ОМ читается «основная морфема».) Ср. *вагоностроитель — вагоностроительный, кинематограф — кинематографический, красноречивый — красноречивость.*

0704. Для $[(aBckdEf|\phi_i|]\vdash[(aBckdEy)|\phi_j|]$ при $y \neq f$, $y \neq 0$ и $f \neq 0$. Легко показать, что при $y = 0$ или $f = 0$ описываемый случай превращается в предыдущий. Замена одного суффикса алломорфом недопустима. Это относится к *-ущ-, -ющ-, -ащ-, -ящ-, -ом-, -ем-, -им-, -ым-* и др. (см. соглашение об исследуемом материале). Интерпретация: «В языке L для сложного слова $\mathfrak{A}$ существует такое сложное слово $\mathfrak{B}$, что при тождестве частей *aBckdE* основные морфемы отличаются взаимозаменяющими суффиксами f и y или их частями. Иными словами, основные морфемы сложных слов тождественны по модулю (f, y)». Ср. *законодательский — законодательство, пятилетка — пятилетний.*

0705.[18] Для $[(aBckdEf)\,|\,\phi_i\,|]\vdash[(dEf)\,|\,\phi_j\,|]$, где части dEf тождественны, относительно же $|\,\phi_i\,|$ и $|\,\phi_j\,|$ никаких ограничений не накладывается. Интерпретация: «В языке L для сложного слова $\mathfrak{A}$ существует такое несложное, слово $\mathfrak{B}$, что основная морфема слова $\mathfrak{B}$ тождественна последней простой основной морфеме сложного слова $\mathfrak{A}$»[19]. Ср. *вагоностроитель — строитель, острозубый — зуб.*

0706. Для $[(aBckdEf)\,|\,\phi_i\,|]\vdash[(dE)\,|\,\phi_j\,|]$ при $f\neq 0$. Интерпретация: «В языке L для сложного слова $\mathfrak{A}$ существует такое несложное слово $\mathfrak{B}$, что основная морфема слова $\mathfrak{B}$ включается в последнюю простую основную морфему слова $\mathfrak{A}$». Ср. *каменноугольный — уголь, острозаразный — зараза.*

0707. Для $[(aBckdEf)\,|\,\phi_i\,|]\vdash[(dEy)\,|\,\phi_j\,|\,]$ при $f\neq y$, $y\neq 0$ и $f\neq 0$ и условии, что f и y не алломорфы (ср. 0704). Интерпретация: «В языке L для сложного слова $\mathfrak{A}$ существует такое несложное слово $\mathfrak{B}$, что основная морфема слова $\mathfrak{B}$ и последняя простая основная морфема слова $\mathfrak{A}$ тождественны по модулю (f, y)». Ср. *каменноугольный — угольщик, словоразличительный — различать.*

0708. Для $[(aBckdEf)\,|\,\phi_j\,|]\vdash[(dEfy)\,|\,\phi_j\,|]$ при $y\neq 0$. Интерпретация: «В языке L для сложного слова $\mathfrak{A}$ существует такое несложное слово $\mathfrak{B}$, что ОМ$\in\mathfrak{B}$ включает в качестве составной части последнюю простую основную морфему слова $\mathfrak{A}$, при этом они не тождественны». Ср. *ярко-синий — синева, светочувствительный — чувствительность.*

0709. Для $[(aBckdEf)\,|\,\phi_j\,|]\vdash[(aBc)\,|\,\phi_j\,|]$, где части aBc тождественны, а относительно ϕ_i и ϕ_j никаких ограничений не указывается. Интерпретация: «В языке L для сложного слова $\mathfrak{A}$ существует такое несложное слово $\mathfrak{B}$, что ОМ$\in\mathfrak{A}$ включает в качестве составной части основную морфему слова $\mathfrak{B}$, или что основная морфема слова $\mathfrak{B}$ и одна из непоследних основных морфем, простых и сложных, слова $\mathfrak{A}$ тождественны». Ср. *зубоврачебный — зуб, камнедробильный — камень, теплотехник — тепло.*

0710. Для $[(aBckdEf)\,|\,\phi_i\,|]\vdash[(aB)\,|\,\phi_j\,|]$ при $c\neq 0$. Интерпретация: «В языке L для сложного слова $\mathfrak{A}$ существует такое несложное слово $\mathfrak{B}$, что основная морфема слова $\mathfrak{B}$ включается в одну из непоследних основных морфем слова $\mathfrak{A}$». Ср. *серовато-синий — серый, каменноугольный — камень.*

0711. Для $[(aBckdEf)\,|\,\phi_i\,|]\vdash[(aBy)\,|\,\phi_i\,|]$ при $c\neq y$, $y\neq 0$ и $c\neq 0$ и условии, что C и y не являются алломорфами (ср. 0704, 0707). Интерпретация: «В языке L для сложного слова $\mathfrak{A}$ существует такое несложное слово $\mathfrak{B}$, что основная морфема слова $\mathfrak{B}$ и одна из непоследних. основных морфем слова $\mathfrak{A}$ тождественны по модулю (c, y)». Ср. *каменноугольный — каменщик, серовато-синий — серость.*

0712. Для $[(aBckdEf)\,|\phi_i|]-[(aBcy)\,|\,\phi_j\,|]$ при $c\neq y\neq 0$. Интерпретация. «В языке L для сложного слова $\mathfrak{A}$ существует такое несложное слово $\mathfrak{B}$, что основная морфема этого слова $\mathfrak{B}$ включает одну из непоследних основных морфем слова $\mathfrak{A}$, при этом они не тождественны». Ср. *активно-пассивный — активность.*

080. Исходя из соглашений и определений в разделе 3.01, в 080—084 оказывается возможным сформулировать еще два переменных признака, существенных и необходимых для решения поставленной задачи.

081. Усечена ли в сложном слове начальная простая основная морфема? Ср. *технадзор, техминимум, ревтрибунал.*

---

[18] При $r>1$ признаки 0705—0702 проверяются на каждом из r отрезков сложного слова, образующихся в результате (r—1)-кратного отсечения по одной основной простой морфеме слева для признаков 0705—0708 и справа для 0709—0712.

[19] В современной дескриптивной лингвистике этому признаку соответствовал бы следующий признак: «В языке L часть dEf слова $\mathfrak{A}$ может быть свободной формой и входить в слово $\mathfrak{B}$». О понятиях связанной и свободной форм см.: K. Pike. Phonemics. Ann Arbour, 1947, стр. 162—163, 234, 238.

082. Усечена ли в сложном слове конечная простая основная морфема? Ср. *ревком*, *райком*, *рыбхоз*, *универмаг*, *торгпред*.

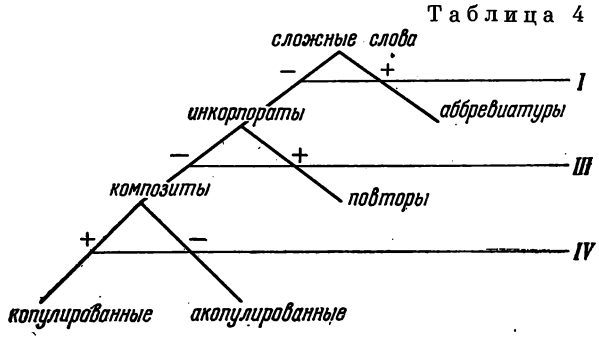
09. Выделенные признаки можно суммировать в табл. 3.

Таблица 3

Номер раздела	Номер по порядку	Формулировка признака
081	I	Усечена ли начальная основная морфема?
082	IA	Усечена ли конечная основная морфема?
0412	II	$r = 1$ или $r > 1$
0413	III	$aBc \subseteq dEf$ или $\overline{aBc \subseteq dEf}$ ($dEf \subseteq aBc$ или $\overline{dEf \subseteq aBc}$)
0412	IV	$k = 0$ или $k = 1$
1701	1	Для $[(aBckdEf(\ /\phi_i/]$ $\vdash$ $[(aBckdEf)\ /\phi_i/]$
1702	2	Для $[(aBckdEf)\ /\phi_i/]$ $\vdash$ $[(aBckdE)\ /\phi_i/]$
1703	3	Для $[(aBckdEf)\ /\phi_i/]$ $\vdash$ $[(aBckdEfy)\ /\phi_i/]$
1704	4	Для $[(aBckdEf)\ /\phi_i/]$ $\vdash$ $[(aBckdEy)\ /\phi_i/]$
1705	5	Для $[(aBckdEf)\ /\phi_i/]$ $\vdash$ $[(dEf)\ /\phi_i/]$
1706	6	Для $[(aBckdEf)\ /\phi_i/]$ $\vdash$ $[(dE)\ /\phi_i/]$
1707	7	Для $[(aBckdEf)\ /\phi_i/]$ $\vdash$ $[(dEy)\ /\phi_i/]$
1708	8	Для $[(aBckdEf)\ /\phi_i/]$ $\vdash$ $[(dEfy)\ /\phi_i/]$
1709	9	Для $[(aBckdEf)\ /\phi_i/]$ $\vdash$ $[(aBc)\ /\phi_i/]$
0710	10	Для $[(aBckdEf)\ /\phi_i/]$ $\vdash$ $[(aB)\ /\phi_i/]$
0711	11	Для $[(aBckdEf)\ /\phi_i/]$ $\vdash$ $[(aBy)\ /\phi_i/]$
0712	12	Для $[(aBckdEf)\ /\phi_i/]$ $\vdash$ $[(aBcy)\ /\phi_i/]$

10. Применив к определению сложного слова операторы проверки признаков I, III, IV, получим рабочую классификацию сложных слов, каждый элемент которой определяется набором признаков. Набор признаков для каждой категории определяется восхождением от соответствующего узла дерева вверх к вершине «Сложные слова» (см. табл. 4).

Таблица 4



Наборы признаков для определения каждой из категорий сложных слов могут быть представлены следующим образом:

Категория	I	III	IV
Аббревиатуры	+		
Инкорпораторы	—		
Повторы	—	+	
Композиты	—	—	
Композиты, копулированные	—	—	+
Композиты, акопулированные	—	—	—

3. СОГЛАШЕНИЯ

01. Соответствие (+) или несоответствие (—) одного из наблюдаемых признаков некоторого слова одному из признаков, перечисленных в табл. 3, назовем значением этого признака слова.

Установление значения признака, запись значения в матрицу сложного слова вида

Номера сложных слов	Признаки						
	1	2	3	4	5	. . .	n

сведение тождественных матриц первого вида в обобщенную матрицу вида

Номера матриц	Номера вошедших сложных слов	Общее количество сложных слов	Признаки					
			1	2	3	4	. . .	n

предполагается проведенным на предварительном этапе исследования.

В результате этого предварительного этапа на дальнейшую обработку поступает совокупность матриц второго вида.

02. Каждая такая матрица могла бы быть названа типом сложного слова. Анализ сложных слов далее сводится к анализу таких матриц.

03. Если основу систематизации как процедуры разбиения на классы составляет поиск и установление соотношений данного описываемого и n переменных признаков, то число классов C, возможное при одной из последовательностей n признаков, составит величину $C=2^n$. Если число варьирующихся последовательностей для n признаков составляет величину $n!$, то теоретически мыслимо $C=n!\ 2^n$ классов. Для нашего случая эта величина составляет $17!2^{17}$.

Уменьшение числа возможных «дорожек» может быть осуществлено за счет выбора одной из 17! последовательностей признаков (в данном случае она выбрана интуитивно) и как результат применения такой процедуры, которая привела бы к минимизации числа классов. Подобная процедура излагается в разделе 4.

04. Логическая простота процедур, наличие формализованной словарной информации при каждом слове и в обобщенной матрице сложного слова делают возможным постановку этой задачи на электронно-счетных машинах. Обилие материала и достаточно большое число признаков, делающие ручной перебор немыслимым, вызывают необходимость механизации этого этапа исследования. Задача лингвиста свелась бы тогда к содержательной интерпретации полученных результатов. Необходимость и возможность механизации анализа подобного типа подкрепляется и тем, что в иных областях науки эта задача поставлена и решается.

05. Совокупность типов сложных слов подвергается классификации по описываемому признаку, установлению переменного признака на подмножестве со значением соответствия описываемому признаку, разбиению всего множества типов на основе выделенного переменного признака, установлению однородности полученных подмножеств в отношении описываемого признака. Множество называется однородным, если все его элементы обладают тем же самым значением описываемого признака.

Таким образом, однородность понимается не абсолютно, но относительно данного описываемого признака.

Если подмножества однородны, то дальнейший анализ не производится. При неоднородности подмножеств для каждого из них вся процедура повторяется.

Описанная процедура состоит в разбиении множества типов сложных слов на классы в зависимости от характера отношений между описываемым и переменными признаками. Кроме того, на основе записи проверяемых переменных признаков устанавливается последовательность поиска необходимой информации, т. е. автоматически, человеком или машиной, выписывается алгоритм выявления соотношений описываемого и переменного признаков. В выписанном таким образом алгоритме число шагов, зависящее от числа используемых операторов, будет равно или меньше числа признаков, приписываемых получающимся предельным классам. Это является следствием нерелевантности некоторых признаков для выделения обладающих ими анализируемых типов в классы по данному описываемому признаку, что объясняется отсутствием классов, обладающих противоположным значением пропущенных признаков. Естественным было бы предположить, что в силу избыточности эти нерелевантные признаки приводят к стабильности и устойчивости всей комбинации признаков.

061. К алгоритму предъявлялись следующие требования:

1) Алгоритм должен быть достаточно общим, он должен в большей своей части действовать на любом подмножестве исследуемой совокупности слов (принцип общности). Иной стороной принципа общности является принцип полноты описания: число предельных классов (см. ниже), не являющихся однородными, должно быть меньше числа предельных однородных классов.

2) Число шагов в алгоритме, не приводящих к выделению предельных классов (см. ниже), должно быть минимальным (принцип простоты или экономии).

3) Признаки, положенные в основу алгоритма, а) должны иметь универсальный характер, обеспечивая использование результатов для сравнительно-типологического анализа языков (принцип универсальности), б) должны быть логически связанными на основе некоторых операций их получения, что обеспечивает экономичность и простоту (см. 2), формальную однородность описания и отвлеченность от семантических соображений принцип гомогенности) [20].

062. Легко заметить, что предъявленные требования не являются жесткими. Для каждого из них могла бы быть предложена достаточно простая процедура оценки алгоритма в абсолютных или относительных величинах. Для каждой из таких оценок в принципе можно было бы указать и пределы варьирования этой оценки [21].

IV

А. 01. Ниже предлагается рабочий вариант алгоритма соотнесения признаков, характеризующих морфемный состав сложных слов, и описываемого признака «одно- и неодносоставность просодемоида».

02. В схеме (табл. 5) приводится древовидное изображение алгоритма.

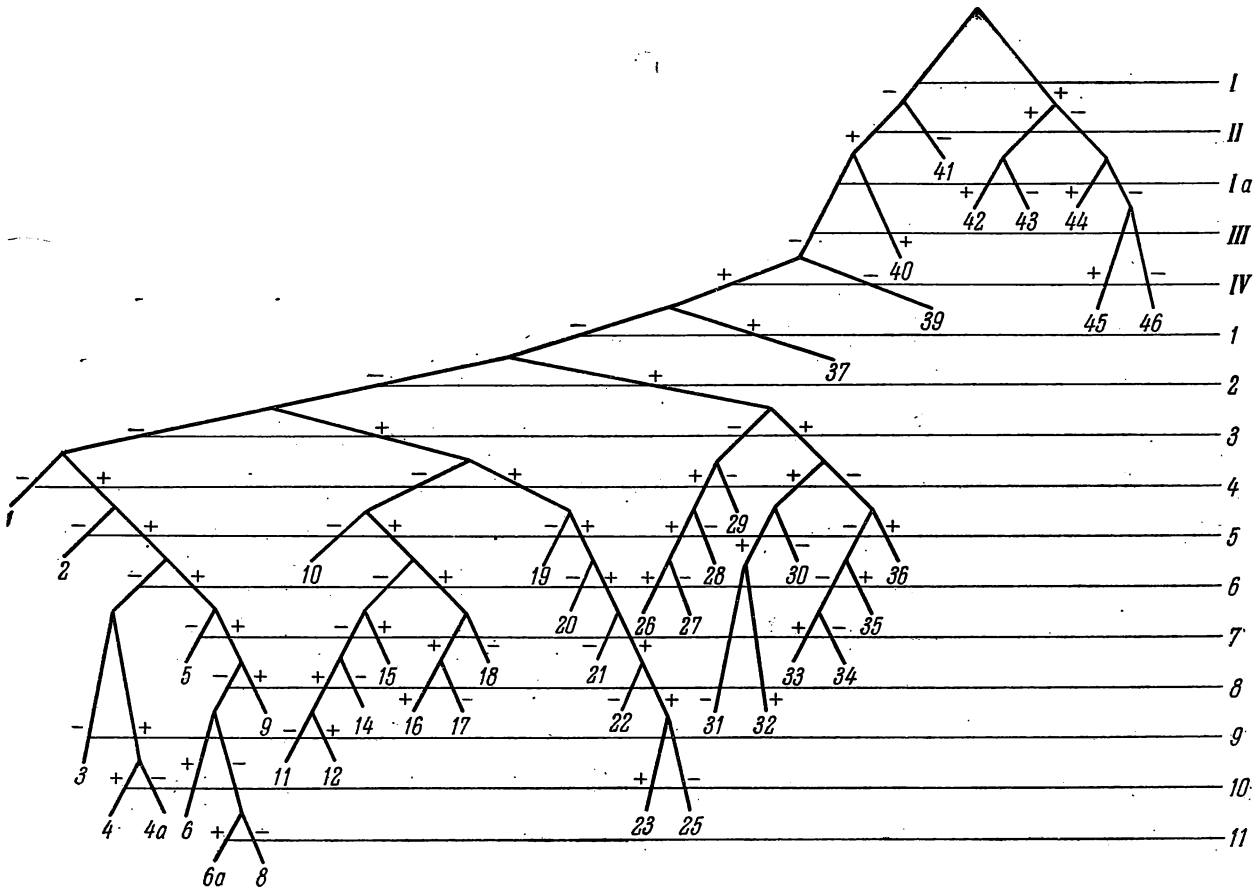
---

[20] О принципе гомогенности см.: С. К. Шаумян. О сущности структурной лингвистики. ВЯ, 1956, № 5, стр. 44: «Всякая теория, объясняя одни факты другими, должна оперировать только такими фактами, которые носят гомогенный характер. С точки зрения ее предмета, в рамках данной теории научное объяснение не может строиться на фактах, лежащих за пределами ее предмета».

Выдвинутые принципы простоты и общности являются логическими следствиями из фундаментальных эпистемологических критериев консеквенции и унификации.

[21] Ср. подобную систему оценок в работе: И. А. Мельчук. Автоматический анализ текстов и некоторые смежные вопросы общей лингвистики. Канд. дисс. М., 1962, разд. III, § 9, стр. 261—277.

Таблица 5


I
II
I a
III
IV
1
2
3
4
5
6
7
8
9
10
11
41
42
43
44
40
39
45
46
37
2
10
19
29
28
30
36
20
26
27
35
5
15
18
21
33
34
9
14
16
17
22
31
32
3
11
12
23
25
4
4a
6
6a
8

В табл. 6 приводится описание процедуры установления связи между значением описываемого признака и набором значений переменных признаков.

Таблица 6

Номер проверяемого признака	Номер следующей операции при реакции +	Номер следующей операции при реакции —
I	II	II
IA	$V_{+2+II+Ia} \vee_{+I-II-Ia}$	$V_{+I+II-Ia} \vee IV_{+I-II-Ia}$
II	$Ia_{+I} \vee III_{-I}$	$Ia_{-I} \vee V$
III	V	IV
IV	$V_{+I-II-Ia} \vee 1_{-I}$	V
1	V	2
2	3	3
3	4	4
4	5	$V_{-2-3+2-3} \vee 5$
5	$8_{+2+3+4} \vee V_{+2+3-4} \vee 6$	$6_{+2+3-4} \vee V$
6	$V_{+2+3-4-5} \vee_{+2} \vee 7_{-2+3} \vee_{-2-3}$	$9_{-1-2-3+4+5} \vee 7_{+2+3} \vee_{-2+3-4} \vee V$
7	$V_{+2} \vee_{-2+3-4+5-6} \vee 8$	$8_{-2+3-4+5-6} \vee V$
8	$9_{-2+3-4+5-6} \vee 10_{-2+3+4}$	$10_{-2-3+4} \vee V$
9	$10_{-2-3+4} \vee V_{-2+3}$	V
10	V	$11_{-2-3+4+6} \vee V$
11	V	V
V	По табл. 8 и набором значений признаков установи номера классов. Если номера классов 22, 29, 36, 37, см. VI, если нет — см. VII	
VI	Проверь, оканчиваются ли слова на *-ый(-ий)*, *-ое(-ее)*, *-ая(-яя)*. Если да, припиши к номеру класса нижний индекс *n*, если нет, припиши нижний индекс $\bar{n}$. См. VII	
VII	По номеру класса в графах 1, 2, 3 табл. 9 установи характер связи между переменными и описываемыми признаками	

03. Операторы I—IV, 1—12 рассчитаны на простую проверку переменных признаков, перечисленных в табл. 3. Таким же является и оператор VI. Операторы V и VII являются операторами конъюнктивной проверки.

Б. Если исходные признаки, вытекающие из определения сложного слова, являются относительно универсальными [22], так как опираются на понятие морфемы — единицы универсального характера, тогда сочетание значений этих исходных, элементарных признаков образует составной производный признак, характерный только для данного языка или только для данной группы языков. Из обилия возможных классов (2^{17}) реализуется всего лишь 42. Подобный подход обеспечивает сравнительно-типологическое изучение языков на основе количественного

[22] Ср. P. Garvin. Указ. соч., стр. 368.

Таблица 7

Номера классов	Переменные признаки																	Примеры
	I	Ia	II	III	IV	1	2	3	4	5	6	7	8	9	10	11	12	
1	—	—	+	—	+	—	—	—	—									*рудо-желтый, садово-парковый, язычно-нёбные*
2	—	—	+	—	+	—	—	—	+	—								*раболепствовать, разноязычие, рукомойник, резьбонакатные*
3	—	—	+	—	+	—	—	—	+	+	—	—	+	—				*умиротворение, ублаготворить*
4	—	—	+	—	+	—	—	—	+	+	—	—	—	+	—			*паенакопление, рабовладение, равностоящий, рукополагать*
4а	—	—	+	—	+	—	—	—	+	+	—	—	—	+	+			*разночтение*
5	—	—	+	—	+	—	—	—	+	+	+	—						*пластинчато-жаберные, сердечно-сосудистый, универсально-фрезерный*
6	—	—	+	—	+	—	—	—	+	+	+	+	—	+	+			*ткацко-отделочный, фуговально-распиловочный, резьбо-прорезной*
6а	—	—	+	—	+	—	—	—	+	+	+	+	—	+	—	+		*початкоцветение*
8	—	—	+	—	+	—	—	—	+	+	+	+	—	+	—	—		*шелкообрабатывающий, шелкотканный, товарообращение, разномыслящий, рукоположение, рукоплескать*
9	—	—	+	—	+	—	—	—	+	+	+	+	+					*сосудисто-волокнистый, рыбопропускной, трубопроводчик*
10	—	—	+	—	+	—	—	+	—	—								*табаковод, уролог, рыжекудрый, толстоносый*
11	—	—	+	—	+	—	—	+	—	+	—	—	+	—				*реверберометр, топограф, тропосфера*
12	—	—	+	—	+	—	—	+	—	+	—	—	+	+				*паропровод, пенобетон, тепловоз, тигролов, раздельнополый*
14	—	—	+	—	+	—	—	+	—	+	—	—	—					*разнобой, тяжеловооруженный*
15	—	—	+	—	+	—	—	+	—	+	—	+						*повторноводящий, трудновоспитуемый*
16	—	—	+	—	+	—	—	+	—	+	+	+	+					*труднодоступный, труднопроходимый, платежеспособный,*
17	—	—	+	—	+	—	—	+	—	+	+	+	—					*турбогенераторный*
18	—	—	+	—	+	—	—	+	—	+	+	—						*узкопленочный*
19	—	—	+	—	+	—	—	+	+	—								*разностопный, ратоборец*
20	—	—	+	—	+	—	—	+	+	+	—							*почвоуглубитель, рассадосажалка*
21	—	—	+	—	+	—	—	+	+	+	+	—						*рудоподъемный*
22	—	—	+	—	+	—	—	+	+	+	+	+	—					*престолонаследник, рыботорговец, тазобедренный, тысячелетний, тонкосуконный*
23	—	—	+	—	+	—	—	+	+	+	+	+	+	+	+			*санитарно-дезинфекционный, торгово-промышленный*
25	—	—	+	—	+	—	—	+	+	+	+	+	+	+	—			*углепромышленник, тонковолокнистый, теплопередача, тугоплавкий*

Таблица 7 (*окончание*)

Номера классов	Переменные признаки																	Примеры
	I	Ia	II	III	IV	1	2	3	4	5	6	7	8	9	10	11	12	
26	—	—	+	—	+	—	+	—	+	+	+							*плодоовощной, полезащитный, шелкопрядение, таксомоторный, турбобурение*
27	—	—	+	—	+	—	+	—	+	+	—							*расоведение; рукописание*
28	—	—	+	—	+	—	+	—	+	—								*ратоборствовать, ротозейничать, театроведение, теплопроводность, роголистниковые*
29	—	—	+	—	+	—	+	—	—									*разноцветность, раннеспелость, угро-финский*
30	—	—	+	—	+	—	+	+	+	—								*углеродистый, торпедоносный, типографщик, тутоводство*
31	—	—	+	—	+	—	+	+	+	+	+	+	—					*торфоперегнойный, углепогрузочный, пустынножительство, шелкопрядильщик*
32	—	—	+	—	+	—	+	+	+	+	+	+	+					*рельсопрокатный, рыбопромысловый, рельсотранспортер, рыбоподъемник*
33	—	—	+	—	+	—	+	+	—	—	—	+						*тенелюбивый, трудолюбивый*
34	—	—	+	—	+	—	+	+	—	—	—	—						*рогохвостый*
35	—	—	+	—	+	—	+	+	—	—	+							*тепловозный, тихоходный*
36	—	—	+	—	+	—	+	+	—	+								*рассадопосадочный, рыборазделочный, углекислота*
37	—	—	+	—	+	+												*рогоносец, рыбокоптильня, раболепный, умиротворенный тысячеголовый, улиткообразный*
39	—	—	+	—	—													*погранзастава, рай-дерево, трехверстка, утильсырьевой четырехосный, тихо-смирно, шито-крыто*
40	—	—	+	+														*перво-наперво, раным-рано, чуть-чуть, пьяный-распьяный, синий-синий, чет-нечет, идти-идти*
41	—	—	—															*паровозоремонтный, серо-буро-малиновый, рентгенокимограф, финно-угроведение*
42	+	+	+															*рабкооп, торгпредство, универмаг*
43	+	—	+															*пединститут, педпрактика, ревтрибунал, райсовет, угрозыск*
44	+	+	—															*рабселькор, райисполком, рабселькоровский*
45	+	—	—	0	+													*райпотребсоюз, райсельхозотдел*
46	+	—	—	0	—													*физкультурно-массовый*

Таблица 8

	Исследуемый признак \ Номер класса	1	2	3	4	4а	5	6	6а	8
I	одноударность		+	+		+			+	
II	неодноударность	+					+			
III	одноударность и неодноударность				+			+		+

	9	10	11	12	14	15	16	17	18	19	20	21	22п	22п̄
I		+	+		+				+	+		+		
II	+					+		+					+	
III				+			+				+			+

	23	25	26	27	28	29п	29п̄	30	31	32	33	34
I				+	+		+	+		+		+
II	+					+			+		+	
III		+	+									

	35	36п	36п̄	37п	37п̄	39	40	41	42	43	44	45	46
I	+	+			+				+		+	+	
II			+			+	+	+					+
III				+						+			

Примечания. Знак + обозначает наличие признака. Знак — обозначает отсутствие признака, 0 (табл. 7) — неприменимость оператора, $\vee$ — или ... или.

Знаком дизъюнкции разграничиваются возможные ходы ветвления в алгоритме анализа, указатели следующего шага располагаются по увеличению общности слева направо (см. табл. 6).

Сложный нижний индекс у указателя шага (табл. 6), например, 6_{+2+3-4} (правая колонка, строка 5) предназначен для отыскания всей предшествующей «истории» анализа и определения ветви в древовидном изображении процедуры. Эта запись читается: «Слово, получившее при проверке пятого признака — (минус) и в матрице которого для второго признака + (плюс), для третьего признака + (плюс), но для четвертого признака — (минус), должно быть проверено на шестой признак».

сопоставления полученных одним и тем же способом описаний и измерения их близости. Поскольку эта тема выходит за рамки предмета работы, ограничимся сказанным.

04. Зависимости между признаками могут быть различного характера. Если исследование носит статический характер, то естественно рассмотреть отношения комплементарности, спецификации и автономии [23]. Все эти отношения подразумевают двухсторонний анализ от переменных признаков к описываемому и от описываемого к переменным. В настоящей работе рассматривается только одна сторона этих отношений — отношения между переменными и описываемым признаками. Этому анализу посвящен раздел V.

V

01. Описание языка или его отдельных частей, раскрывающее предсказуемостное соотношение переменных и описываемых признаков, характерных для этого языка или его отдельных частей, и самих этих описываемых признаков, назовем корреляционной системой.

Корреляционные системы в зависимости от характера связи переменных признаков с описываемыми признаками могут быть либо однородными (C_0), либо смешанными (C_c). Однородные корреляционные системы языка могут быть:

1) системами ненаправленной корреляции категорий различных уровней описания (далее и в аналогичных случаях — система. . . корреляции) при генерализованной импликации *Ig* описываемых признаков переменными;

2) системами избирательной, направленной корреляции при привативной импликации *Ip* описываемого признака переменными признаками;

3) системами независимых отношений между описываемым и переменными признаками.

Системы смешанного типа (C_c) включают элементы всех перечисленных выше систем. Система сложных слов русского языка является системой последнего типа, когда она исследуется для выяснения характера признака «односоставность — неодносоставность просодемоида, оформляющего сложное слово».

Следует отметить, что подсистемы такой системы по отношению друг к другу являются не только взаимоисключающими, но и взаимосвязанными и взаимообусловленными в силу логического единства алгоритма, лежащего в основе их выделения. Исходя из сказанного, зависимость описываемого признака от других признаков можно представить в следующей форме:

$$F_i(\mathfrak{A}) = f[Y_i(\mathfrak{A}),\quad Z_i(\mathfrak{A}\mathfrak{B}\mathfrak{L}),\quad X_i(L\mathfrak{A})],$$

где $F_i(\mathfrak{A})$ — дифференцируемые свойства описываемой совокупности $\mathfrak{A}$;

$Y_i(\mathfrak{A})$ — переменные свойства или свойство совокупности $\mathfrak{A}$, представляемые набором переменных признаков;

$Z_i(\mathfrak{A}\mathfrak{B}\mathfrak{L})$ — ограничения, накладываемые на проявление этих свойств и определяемые связью совокупности $\mathfrak{A}$ с совокупностью или совокупностями $\mathfrak{B}$ и $\mathfrak{L}$ языка L;

---

[23] Л. Ельмслев. Пролегомены к теории языка. «Новое в лингвистике». I, М., 1960, стр. 284—285.

$X_i(L\mathfrak{A})$ — ограничения, накладываемые на свойства $Y_i(\mathfrak{A})$, определяемые характером языка L.

02. Подсистема ненаправленной корреляции. Если оператор, направленный на получение информации о значении некоторого элементарного признака, отражающего $Y_i(\mathfrak{A})$, действует так, что по крайней мере один из выделяемых им классов однороден по значению описываемого признака Fd_i, то такой класс назовем предельным классом, а оператор — предельным оператором по отношению к этому классу. Если этот оператор оказывается предельным в любом случае его использования на любом подмножестве $\mathfrak{A}'$ совокупности $\mathfrak{A}$, назовем его общепредельным $On_0(\mathfrak{A})$. В этом случае между значениями переменных и описываемого признаков возможны следующие предсказуемостные отношения:

Переменные признаки	Дифференцируемые признаки
+	+
+	—
—	+
—	—

Совокупность этих отношений между признаками составляет отношение генерализованной импликации описываемого признака $Ig(Fd_i)$, условием которой является однородность предельных классов в отношении дифференцируемого признака.

Предельные классы, выделяемые применением общепредельного оператора, связаны дизъюнктивно, а подмножества исследуемой совокупности — конъюнктивно, т. е. $\mathfrak{A}'_i \,\&\, \mathfrak{A}''_i \& \ldots \& \,\mathfrak{A}^n_i$, где всякий $\mathfrak{A}^k_i \in \mathfrak{A}_i$.

В корреляционной системе сложных слов современного русского языка такими отношениями связаны:

1) односоставность просодемоида и признаки, характеризующие морфемную структуру сложных слов, принадлежащих к классам 2, 3, 4а, 10, 11, 14, 18, 19, 21, 27, 28, 30, 32, 34, 35, 42, 44, 45 (267 слов);

2) неодносоставность просодемоида и признаки, характеризующие морфемную структуру сложных слов, принадлежащих к классам 1, 5, 6а, 9, 15, 17, 20, 23, 31, 33, 39, 40, 41 и 46 (590 слов).

Подсистема направленной корреляции. Всякий предельный оператор, не являющийся общепредельным по отношению к некоторому классу, назовем автономно-предельным $On_a(K_j)$, где $K_j \in \mathfrak{A}_i$.

По определению On_a предален только по отношению к некоторым классам совокупности $\mathfrak{A}_i$, но не ко всей этой совокупности. Иными словами, подмножества описываемой совокупности связаны дизъюнктивно, в формальной записи $\mathfrak{A}'_i \vee \mathfrak{A}''_i \vee \ldots \vee \mathfrak{A}^n_i$, где всякий $\mathfrak{A}^k_i \in \mathfrak{A}_i$.

В этом случае будем говорить, что между переменным признаком Fp'_f и описываемым признаком Fd_i имеет место импликация Ip класса $K_j \in \mathfrak{A}_i$, которую назовем привативной. Отношение Ip включает те же отношения признаков, что и Ig.

В корреляционной системе сложных слов современного русского языка отношениями Ip связаны «односоставность — неодносоставность» просодемоида с признаками, характеризующими морфемную структуру сложных слов, принадлежащих к непересекающимся подклассам классов 29, 36, 37 и 22 (всего 29 слов).

Разумным было бы, видимо, считать, что свойства сложных слов, определяемые признаками, связь между которыми характеризуется отношениями Ig, являются относительно инвариантными или избирательно инвариантными в противовес тем случаям, где как следствие

отношений генерализованной импликации между признаками свойства объектов общеинвариантны. Но этот вопрос выходит за пределы темы работы и требует особого рассмотрения.

Подсистема независимых отношений. Если ни один из предельных операторов и ни одно из их объединений не приводит к выделению предельного класса объектов, то по отношению ко всей данной совокупности объектов переменные и описываемые признаки связаны отношениями независимости. В корреляционной системе сложных слов современного русского языка эти отношения характерны для признаков, определяющих сложные слова, принадлежащие к классам 4, 6, 8, 12, 16, 22, 25, 26, 37, 43. Как число классов, так и мощность всей подсистемы оказались близкими к ожидавшимся величинам.

Отношение независимости может быть вызвано:

1. Ограничениями, накладываемыми на исследование соглашением об аспектах исследования. Можно предполагать, что отношения дополнительности статического и динамического аспектов анализа выражаются также и в том, что всегда будет остаток — подсистемы независимых отношений, которые описываются только при одном из дополняющих друг друга аспектов анализа.

Все иные источники независимых отношений характерны только для тех элементов корреляционной системы сложных слов, которые остаются неописанными как при статическом, так и динамическом изучении совокупности.

2. Недостаточностью системы используемых элементарных признаков, что в целях компенсации приводит к необходимости введения новых признаков того же самого уровня описания (фонологического, морфологического или синтаксического).

3. Возмущениями, вносимыми в структуру языка структурой речи.

4. Случайными внеструктурными возмущениями (например заимствованиями и процессом ассимиляции заимствований). Подобные случаи исключены по соглашению об анализируемом материале.

5. Возмущениями, оказываемыми системами других уровней языка[24].

Из всех перечисленных источников возмущений этот последний имеет наиболее серьезное значение, как это легко можно было бы показать, проанализировав, например, класс 37 и рассмотрев его долю в системе независимых отношений. С теоретической и практической точки зрения это требует последовательного различения ступени конструктов и ступени наблюдения.

В данной работе анализ отношений независимости не проводится, так как, требуя последовательного рассмотрения возможных типов возмущений, он является особой задачей. Обязательность этого анализа вытекает хотя бы из необходимости разграничения $Z_i(\mathfrak{AB}\mathfrak{C})$ и $X_i(L\mathfrak{A})$, требуемого при построении синтезирующей грамматики.

03. Признак «одно- и неодносоставность просодемоида, оформляющего сложное слово» исходит из содержания лингвистических категорий просодемы и просодемоида[25]. Некоторые разъяснения этих терминов были даны выше (раздел I, § 2). Проведенный анализ позволяет сделать некоторые замечания о специфике связи просодемы и просодемоида.

---

[24] Например, количеством слогов от начала слова до последнего ударного слога в этом слове включительно. См. З. П. Донова. К вопросу об ударении в сложных словах. «Ученые записки Орехово-Зуевского пед. ин-та», т. II, вып. 1, 1955.

[25] С. К. Шаумян. Проблемы теоретической фонологии, стр. 163.

В нашем анализе уровень грамматики представлен переменными признаками, которые характеризуют морфемный состав сложного и сложносокращенного слова. Часть последовательностей значений этих признаков диагностирует один просодемоид, другие — устойчивую комбинацию из нескольких просодемоидов. Такую устойчивую комбинацию просодемоидов, подобно воплощению просодемы в первом случае, разумно, видимо, рассматривать как некоторую целостную лингвистическую единицу.

Во-первых, она представляет собой воплощение просодемы, оформляющей слово как семиотическую единицу[26], поэтому она не может не быть целостна, едина.

Во-вторых, единство устойчивой комбинации просодемоидов, присущих слову, вытекает из следующего рассуждения. Пусть просодемоид всегда односоставен. Поскольку наличие просодемоида определяется некоторой последовательностью значений переменных признаков, то, если в слове два просодемоида, независимых по отношению друг к другу, то должно быть и две нетождественные последовательности значений переменных признаков. Анализ противоречит этому выводу, логически следующему из нашей посылки. Значит, она ложна. Тогда естественен такой вывод: так как последовательность значений признаков, диагностирующая два или более просодемоида, одна и только одна, а никакой другой последовательности значений этих же признаков для предсказания этих просодемоидов не существует, то в силу единственности такой последовательности оба просодемоида едины и служат воплощением одной единственной просодемы. Иными словами, существование одного из них непременно определяет и существование других даже в том случае, если они дистантны.

04. Сказанное может быть обобщено в виде тезисов, требующих дальнейшего рассмотрения.

041. На ступени конструктов каждое слово может иметь одну и только одну просодему, характер которой определяется особенностями связи уровня фонологии и уровня грамматики и спецификой отношений между ступенью конструктов и ступенью наблюдения в фонологии.

042. На ступени наблюдения, субстанциональном уровне, уровне непосредственного анализа речи, просодеме соответствует просодемоид, односоставный или неодносоставный. Специфика просодемоида как лингвистической единицы определяется характером просодемы, спецификой отношений между ступенями конструктов и наблюдения и особенностями физических и антропофонических характеристик речи.

043. Одной просодеме соответствует один и только один просодемоид, одно- или неодносоставный.

044. Специфика связи между ступенями конструктов и наблюдения в фонологии, в частности при рассмотрении просодических элементов, определяется особенностями последовательностей значений признаков, конституирующих категорий противоположного плана, в частности формы плана выражения.

045. Связь между лингвистическими категориями ступени конструктов и ступени наблюдения характеризуется отношениями гомоморфизма.

05. Некоторые методические и методологические замечания. Недостаточность методов экспериментального исследования просодических элементов и отсутствие иных объективных методов их изучения требуют

---

[26] «Каждому самостоятельному слову обычно соответствует одно ударение» (Р. И. Аванесов. Кратчайшая звуковая единица в составе слова и морфемы. «Вопросы грамматического строя». М., 1955, стр. 121). Исходной гипотезой в этом случае является утверждение: «Слово, в противоположность иным средствам номинации, оформляется одной и только одной просодемой».

определения качества просодемоида и его специфики (подвижность — неподвижность и т. д.) в разных случаях на основе формального исследования связи фонологии и грамматики. В этом случае мы исходим из соотнесения результатов анализа на ступени наблюдения и ступени конструктов формы плана содержания, а от них — к единицам ступени наблюдения формы плана выражения на основе предсказуемости последних соответствующими элементами плана содержания. Тогда с помощью процедур постулирования конструктов можно перейти к соответствующим конструктам плана выражения.

Предложенный анализ некоторых случаев неодносоставности просодемоидов представляет собою образец подобной методики (табл. 9).

При изучении конструктов уровня грамматики представляется в ряде случаев возможным на основе фактов предсказуемостного анализа восхождение от данных корреспонденции результатов анализа на ступени наблюдения и ступени конструктов уровня фонологии, к единицам ступени наблюдения и конструктам плана содержания. Сказанное имплицитно содержит требование исходить при характеристике лингвистических сущностей из данных ступени наблюдения как плана выражения, так и плана содержания, при этом несущественно, к какому плану относится изучаемое явление. К сожалению, по отношению к ударению по указанным выше причинам достаточно целесообразное использование этой процедуры представляется невозможным.

Таблица 9

	Форма плана выражения	Форма плана содержания		
	Просодема	Морфема Сложное слово Признаки сложного слова		Ступень конструктов
Процедура постулирования конструктов	Просодемоид Признаки просодемоида анализируемого сложного слова	Признаки анализируемого сложного слова	Алгоритм анализа	Ступень наблюдения
	↑ Алгоритм предсказания → ↓			

ON SOME CASES OF NON-UNITARINESS OF PROSODEMOIDS

An experiment of operational analysis

Summary

The paper deals with an operational method of predicting unitary and non-unitary prosodemoids — an explicata of the «accent» according to S. K. Šaumjan' s two-level theory of phonology. The first antinomy of accentual word structure observed by E. D. Polivanov is proved to be eliminated by introducing notions of prosodemoids and prosodema.

An algorithm of predicting unitariness or non-unitariness of prosodemoid consists mainly in checking up morphological characteristics of the compound. A set of these is developed on the basis of the recursive definition of the compound, which may be expressed by the following formula: $[(dEfk)^r (aBc)]|\phi|$ where a, d — prefixes, c, f — suffixes, B, E — roots, k — a conjunctional morpheme, (ϕ) — inflection, r — the number of non-final elementary stems which the compound stem consists of.

An interpretation of the abbreviated compound is suggested. Accordingly a classification of compounds is proposed. The procedure of automatic developing algorithms for predicting some characteristics on the basis of others is described.

Э. Василиу

СТРУКТУРА КОНСОНАНТНЫХ ГРУПП В РУМЫНСКОМ ЯЗЫКЕ

0. Цель настоящей статьи — дать описание консонантных групп румынского языка с точки зрения структуры слога.

Статья состоит из двух частей. В первой части строится общая теория консонантных групп и излагается метасистема, с помощью которой будет проводиться описание конкретных структур. В данном случае структура консонантных групп румынского языка может посредством этой метасистемы сопоставляться со структурой консонантных групп в других языках. Во второй части дается описание структуры консонантных групп румынского языка на базе метасистемы, построенной в первой части.

1. Общая теория

В качестве первичных элементов принимаются следующие: категория гласных (V), категория согласных (C), пауза ($\#$).

Определение 1. Мы говорим, что консонантный сегмент SC, который может встречаться между паузой и гласным, является начальным сегментом (SI). Это можно выразить следующим образом:

(1) $(SC)\ (\# SCV \rightarrow SC \in SI) =_{Df} SI$.

Определение 1а. Согласный, принадлежащий к классу SI, называется начальным согласным (i):

(1а) $C \in SI =_{Df} i$.

Определение 2. Консонантный сегмент SC, который может встречаться между гласным и паузой, является конечным сегментом (SF):

(2) $(SC)\ (VSC \# \rightarrow SC \in SF) =_{Df} SF$.

Определение 2а. Согласный, принадлежащий к сегменту SF, называется конечным согласным (f):

(2а) $C \in SF =_{Df} f$.

Гипотеза *а*. Начальный сегмент может иметь вид i, ii. или iii, но никогда не содержит более трех согласных:

$$(a)\ SI = (i \vee ii \vee iii) \cdot \sim iii \ldots i_n.$$

Гипотеза *б*. Конечный сегмент может иметь вид f, ff или fff, но никогда не содержит более трех согласных:

$$(б)\ SF = (f \vee ff \vee fff) \cdot \sim fff \ldots f_n.$$

Из этого следует, что, например,

$$ffff = \sim SF.$$

Обозначим через x, y и z три различных согласных. Скобки $\langle\rangle$ означают, что элементы, заключенные в эти скобки, могут встречаться рядом на синтагматической оси.

Например, $\langle xy\rangle$ означает, что x встречается рядом с y.

1. 1. Дистрибутивные классы

Определение 3. Определим класс I_1 посленачальных согласных следующим образом:

(3) $(y)\ [\langle xy\rangle \to (xy\ \ V \cdot \sim yx\ V)] =_{Df} y \in I_1$.

В словесной форме это определение гласит: если для каждого y при сочетании x и y члены сочетания встречаются в порядке xyV, но никогда не встречаются в порядке yxV, то y по определению принадлежит классу I_1 посленачальных согласных.

Определение 3а. Согласный i, принадлежащий к классу I_1, есть посленачальный согласный (i_1):

(3а) $y \in I_1 =_{Df} i_1$.

Определение 4. Определим класс I_2 собственно начальных согласных следующим образом:

(4) $(x)\ [\langle xi_1\rangle \to (x_1\ V \cdot \sim i_1\ xV)] =_{Df} x \in I_2$.

Если для каждого x при сочетании x с согласным i_1 члены сочетания встречаются всегда в порядке xi_1V и никогда не встречаются в порядке i_1xV, то x, по определению, принадлежит классу I_2 собственно начальных согласных.

Определение 4а. Согласный i, принадлежащий к классу I_2, есть собственно начальный согласный:

(4а) $x \in I_2 =_{Df} i_2$.

Определение 5. Определим класс I_3 предначальных согласных следующим образом:

(5) $(z)\ \{\langle z_1i_2i_1\rangle \to [zi_2i_1 \cdot \sim (i_2zi_1V \vee i_1zi_2V \vee i_1i_2zV \vee i_2i_1zV]\} =_{Df} z \in I_3$.

Если для каждого z при сочетании z с i_1 и i_2 члены сочетания встречаются в порядке zi_2i_1V и никогда не встречаются в порядке i_2zi_1V, zi_2V, i_1i_2zV или i_2i_1zV, то z, по определению, принадлежит классу I_3 предначальных согласных.

Определение 5а. Согласный i, принадлежащий к классу I_3, есть предначальный согласный (i_3):

(5а) $z \in I_3 =_{Df} i_3$.

Определение 6. Определим класс F_1 предконечных согласных следующим образом:

(6) $(y)\ [\langle x_1y\rangle \to (Vyx \cdot \sim Vxy)] =_{Df} y \in F_1$.

Если для каждого y при сочетании x с y члены сочетания встречаются в порядке Vyx и никогда не встречаются в порядке Vxy, то y, по определению, принадлежит классу F_1 предконечных согласных.

Определение 6а. Согласный f, принадлежащий к классу F_1, есть предконечный согласный (f_1):

(6а) $y \in F_1 =_{Df} f_1$.

Определение 7. Определим класс F_2 собственно конечных согласных следующим образом:

(7) $(x)\ [\langle xf_1\rangle \to (Vf_1x \cdot \sim Vxf_1)] =_{Df} x \in F_2$.

Если для каждого x при сочетании x и f_2 члены сочетания встречаются в порядке Vf_1x и никогда не встречаются в порядке Vxf_2, то x, по определению, принадлежит классу F_2 собственно конечных согласных.

Определение 7а. Согласный f, принадлежащий к классу F_2, есть собственно конечный согласный (f_2):

(7а) $x \in F_2 =_{Df} f_2$.

Определение 8. Определим класс F_3 послеконечных согласных следующим образом:

(8) $(z)\ \{\langle f_1f_2z\rangle \to [Vf_1f_2z \cdot \sim (Vf_1zf_2 \vee Vf_2zf_1 \vee Vzf_2f_1 \vee Vzf_1f_2)]\} =_{Df} z \in F_3$.

Если для каждого z при сочетании z с f_1 и f_2 члены сочетания встречаются в порядке Vf_1f_2z и никогда не встречаются в порядке Vf_1zf_2 или Vf_2zf_1 или Vzf_1f_2 или Vzf_2f_1, то z, по определению, принадлежит классу F_3 послеконечных согласных.

Определение 8а. Согласный f, принадлежащий к классу F_3, есть послеконечный согласный:

(8а) $z \in F_3 = {}_Df f_3$.

1. 1. 1. Вторичные и третичные функции.

Если в определении (5) выразим z через i_3, получим:

(9) $[\langle i_1, i_2, i_3\rangle \rightarrow (i_3i_2i_1V \cdot \sim (i_2i_3i_1V \vee i_1i_3i_2V \vee i_1i_2i_3V \vee i_2i_1i_3V)$.

Импликация (9) определяет катализованный начальный сегмент.

Аналогичным образом можно определить катализованный конечный сегмент:

(10) $\langle f_1, f_2, f_3\rangle \rightarrow [Vf_1f_2f_3 \cdot \sim (Vf_1f_3f_2 \vee Vf_2f_3f_1 \vee Vf_3f_2i_1 \vee Vf_3f_2f_1)]$.

Иными словами, мы можем сказать, что катализованный сегмент — это такой сегмент максимального протяжения, в котором все согласные принадлежат к разным классам и в котором каждый согласный занимает ту позицию, которая фигурирует в определении соответствующего дистрибутивного класса.

Сегменты, для которых это условие не выполнено, суть некатализованные сегменты.

Из определения (4) вытекает, что $I_2 \rightarrow I_1$, так как определение класса I_2 базируется на определении I_1.

Подобно этому из определения (7) вытекает, что $F_2 \rightarrow F_1$.

Из определения (7) вытекает, что I_3 имплицирует существование классов I_2 и I_1 (которые сами находятся в отношении импликации), так как I_3 определяется через I_2 и I_1:

(11) $I_3 \rightarrow (I_2 \rightarrow I_1)$.

Аналогичным образом

(12) $F_3 \rightarrow (F_2 \rightarrow F_1)$.

Если согласные x и y (принадлежащие соответственно к классам X и Y) встречаются в последовательности xy в начальном сегменте и в последовательности yx в конечном сегменте и если $X \rightarrow Y$, мы говорим, что y есть имплицируемый согласный.

Мы говорим, что согласный выполняет вторичную функцию в тех случаях, когда в некатализованном сегменте этот согласный не занимает позицию, которая фигурирует в определении класса, к которому он принадлежит, но занимает позицию имплицированного согласного катализованного сегмента.

Следовательно, если $x \in I_3$ и x встречается в последовательности xi_1, мы говорим, что x выполняет вторичную функцию собственно начального согласного (обозначим его через x'_3). Если $x \in I_2$ и x встречается в последовательности i_2x, мы говорим, что x выполняет вторичную функцию последовательного согласного и т. п.

Мы говорим, что согласный выполняет третичную функцию, если он не занимает позицию, фигурирующую в определении соответствующего класса, и не занимает позицию имплицированного согласного катализованного сегмента.

Следовательно, если $x \in I_3$ и x встречается в последовательности i_2x, мы говорим, что x выполняет третичную функцию посленачального согласного.

Обозначим через $i_{1'}$, $i_{2'}$, $i_{3'}$, $f_{1'}$, $f_{2'}$, $f_{3'}$ согласные, которые выполняют вторичную функцию, а через $i_{1''}$, $i_{2''}$, $i_{3''}$, $f_{1''}$, $f_{2''}$, $f_{3''}$ — согласные, выполняющие третичную функцию. Равенства типа $ii = i_1i_2$, $ii = i_3i_1$ будут указывать, что перед нами сочетание двух согласных, из которых

первый принадлежит классу I_1, а второй — классу I_2 (первое равенство), или первый принадлежит классу I_1, а второй классу I_3 (второе равенство) и т. п.

1. 1. 2. Перечислим все возможные типы катализованных и некатализованных сочетаний согласных.

1. 1. 2. 1. Катализованные сочетания всегда имеют следующий вид:

a) $(iii = i_3 i_2 i_1) \rightarrow i_3 i_2 i_1$;

b) $(fff = f_1 f_2 f_3) \rightarrow f_1 f_2 f_3$.

1. 1. 2. 2. 1. Некатализованное сочетание может состоять из трех согласных, принадлежащих к одному и тому же классу. В этом случае сочетания могут быть следующих видов:

(1a) $(iii = i_1 i_1 i_1) \rightarrow i_1 i_1 i_1$, (1b) $(fff = f_1 f_1 f_1) \rightarrow f_1 f_1 f_1$,

(2a) $(iii = i_2 i_2 i_2) \rightarrow i_2 i_2 i_2$, (2b) $(fff = f_2 f_2 f_2) \rightarrow f_2 f_2 f_2$,

(3a) $(iii = i_3 i_3 i_3) \rightarrow i_3 i_3 f_3$, (3b) $(fff = f_3 f_3 f_3) \rightarrow f_3 f_3 f_3$.

1. 1. 2. 2. 2. Некатализованное сочетание может состоять из двух согласных, принадлежащих к одному и тому же классу. В этом случае сочетания могут быть следующих видов:

(1a) $(ii = i_1 i_1) \rightarrow i_1 f_1$, (1b) $(ff = f_1 f_1) \rightarrow f_1 f_1$,

(2a) $(ii = i_2 i_2) \rightarrow i_2 i_2$, (2b) $(ff = f_2 f_2) \rightarrow f_2 f_2$,

(3a) $(ii = i_3 i_3) \rightarrow i_3 i_3$, (3b) $(ff = f_3 f_3) \rightarrow f_3 f_3$.

1. 1. 2. 2. 3. Некатализованное сочетание может состоять из двух согласных, принадлежащих к разным классам. В этом случае сочетания могут быть следующих видов:

(1a) $(ii = i_2 i_1) \rightarrow i_2 i_1$, (1b) $(ff = f_1 f_2) \rightarrow f_1 f_2$,

(2a) $(ii = i_3 i_2) \rightarrow i_3 i_2$, (2b) $(ff = f_2 f_3) \rightarrow f_2 f_3$,

(3a) $(ii = i_3 i_1) \rightarrow i_3 i_1$, (3b) $(ff = f_1 f_3) \rightarrow f_1 f_3$,

(4a) $(ii = i_2 i_3) \rightarrow i_2 i_3$, (4b) $(ff = f_3 f_2) \rightarrow f_3 f_2$.

Примечание. В этот перечень не включены сочетания, противоречащие определениям дистрибутивных классов, например, сочетания типа $*i_1 i_2$, противоречащие определениям (1), (2), сочетания типа $*i_1 i_3$, противоречащие определению (1) и т. д.

1. 1. 2. 2. 4. Некатализованное сочетание может состоять из трех согласных, два из которых принадлежат к одному и тому же классу, а один принадлежит к другому классу. В этом случае сочетания могут быть следующих видов:

(1a) $(iii = i_2 i_2 i_1) \rightarrow i_2 i_2 i_1$, (1b) $(fff = f_1 f_2 f_3) \rightarrow f_1 f_2 f_2$,

(2a) $(iii = i_3 i_3 i_1) \rightarrow i_3 i_3 i_1$, (2b) $(fff = f_1 f_3 f_3) \rightarrow f_1 f_3 f_3$,

(3a) $(iii = i_3 i_3 i_2) \rightarrow i_3 i_3 i_2$, (3b) $(fff = f_2 f_3 f_3) \rightarrow f_2 f_3 f_3$,

(4a) $(iii = i_3 i_2 i_2) \rightarrow i_3 i_2 i_2$, (4b) $(fff = f_2 f_2 f_3) \rightarrow f_2 f_2 f_3$,

(5a) $(iii = i_3 i_1 i_1) \rightarrow i_3 i_1 i_1$, (5b) $(fff = f_1 f_1 f_3) \rightarrow f_1 f_1 f_3$,

(6a) $(iii = i_2 i_3 i_3) \rightarrow i_2 i_3 i_3$, (6b) $(fff = f_3 f_3 f_2) \rightarrow f_3 f_3 f_2$,

(7a) $(iii = i_3 i_2 i_3) \rightarrow i_3 i_2 i_3$, (7b) $(fff = f_3 f_2 f_3) \rightarrow f_3 f_2 f_3$,

(8a) $(iii = i_1 i_2 i_1) \rightarrow i_1 i_2 i_1$, (8b) $(fff = f_1 f_2 f_1) \rightarrow f_1 f_2 f_1$,

(9a) $(iii = i_1 i_3 i_1) \rightarrow i_1 i_3 i_1$, (9b) $(fff = f_1 f_3 f_1) \rightarrow f_1 f_3 f_1$,

(10a) $(iii = i_2 i_3 i_2) \rightarrow i_2 i_3 i_2$, (10b) $(fff = f_2 f_3 f_2) \rightarrow f_2 f_3 f_2$,

(11a) $(iii = i_3 i_1 i_3) \rightarrow i_3 i_1 i_3$, (11b) $(fff = f_2 f_1 f_3) \rightarrow f_3 f_1 f_3$.

1. 1. 2. 2. 5. Некатализованный сегмент может состоять из одного согласного. Этот сегмент может быть следующих видов:

$$\begin{array}{ll} \text{(1a)}\ (i = i_1) \to i_1, & \text{(1b)}\ (f = f_1) \to f_1, \\ \text{(2a)}\ (i = i_2) \to i_{2'}, & \text{(2b)}\ (f = f_2) \to f_{2'}, \\ \text{(3a)}\ (i = i_3) \to i_{3''}, & \text{(3b)}\ (f = f_3) \to f_{3''}, \\ \text{(4a)}\ (i = i_a) \to i_a, & \text{(4b)}\ (f = f_a) \to f_a. \end{array}$$

1. 2. Рассмотрим структуру интервокальных сегментов. Мы будем использовать знаки *i*, *ii*, *iii* и *f*, *ff*, *fff* следующим образом:

i может иметь любое из значений, указанных в 1. 1. 2. 2. 5, пп. 1a — 4a;

ii может иметь любое из значений, указанных в 1. 1. 2. 2. 2, пп. 1a—3a и в 1. 1. 2. 2. 3, пп. 1a—4a;

iii может иметь любое из значений, указанных в 1. 1. 2. 1, пункт a и b 1. 1. 2. 2. 4, пп. 1a—11a;

f может иметь любое из значений, указанных в 1. 1. 2. 2. 5, пп. 1b—4b;

ff может иметь любое из значений, указанных в 1. 1. 2. 2. 2, пп. 1b—3b и в 1. 1. 2. 2. 3, пп. 1b—4b;

fff может иметь любое из значений, указанных в 1. 1. 2. 1, пункт b или в 1. 1. 2. 2. 4, пп. 1b—11b.

Будем считать, что интервокальное сочетание согласных не может содержать более шести согласных. Иными словами, интервокальное сочетание максимального протяжения имеет вид *VCCCCCCV*.

Определим слогораздел следующим образом:

$$\text{(13)}\ I - I =_{Df} \left[V \left\{ \begin{matrix} C \\ C\ C \\ \text{ø} \end{matrix} \right\} \left\{ \begin{matrix} C \\ \text{ø} \end{matrix} \right\} C \left\{ \begin{matrix} C \\ C\ C \\ \text{ø} \end{matrix} \right\} V = VSFSIV \right] \equiv$$

$$\equiv V \left\{ \begin{matrix} C \\ C\ C \\ \text{ø} \end{matrix} \right\} \left\{ \begin{matrix} C \\ \text{ø} \end{matrix} \right\} - C \left\{ \begin{matrix} C \\ C\ C \\ \text{ø} \end{matrix} \right\} V.$$

В словесной форме: сегмент вида $V \left\{ \begin{matrix} C \\ C\ C \\ \text{ø} \end{matrix} \right\} \left\{ \begin{matrix} C \\ \text{ø} \end{matrix} \right\} - C \left\{ \begin{matrix} C \\ C\ C \\ \text{ø} \end{matrix} \right\} V$,

равный сегменту вида *VSFSIV*, эквивалентен интервокальному консонантному сегменту, внутри которого слогораздел (—) проходит следующим образом:

$$V \left\{ \begin{matrix} C \\ C\ C \\ \text{ø} \end{matrix} \right\} \left\{ \begin{matrix} C \\ \text{ø} \end{matrix} \right\} - C \left\{ \begin{matrix} C \\ C\ C \\ \text{ø} \end{matrix} \right\} V.$$

Сформулируем следующий постулат:

$$\text{(14)} \left[V \left\{ \begin{matrix} C \\ C\ C \\ \text{ø} \end{matrix} \right\} C = V \left\{ \begin{matrix} f \\ f\ f \\ \text{ø} \end{matrix} \right\} fV \equiv V \left\{ \begin{matrix} C \\ C\ C \\ \text{ø} \end{matrix} \right\} - CV.$$

Постулат (14) означает, что в нашей системе слогораздел между согласным и гласным не допускается.

Придав значение «ложь» первому члену эквивалентности (13), мы получим определение интерлюдии[1], т. е. интервокального сегмента,

---

[1] Ср. термин «interlude» у Хокетта (Ch. Hockett, Manual of phonology. Baltimore, 1955).

внутри которого не проходит слогораздел:

$$(15)\ \left[V\left\{\begin{matrix}C\\C\ C\\ø\end{matrix}\right\}\left\{\begin{matrix}C\\ø\end{matrix}\right\}V \neq VSFSIV\right] \equiv V\left\{\begin{matrix}C\\C\ C\\ø\end{matrix}\right\}\left\{\begin{matrix}C\\ø\end{matrix}\right\} \neq C\left\{\begin{matrix}C\\C\ C\\ø\end{matrix}\right\}V =$$
$$= {}_{Df}\left\{\begin{matrix}C\\C\ C\\ø\end{matrix}\right\}\left\{\begin{matrix}C\\ø\end{matrix}\right\}C\left\{\begin{matrix}C\\C\ C\\ø\end{matrix}\right\}.$$

В словесной форме: сегмент вида $V\left\{\begin{matrix}C\\C\ C\\ø\end{matrix}\right\}\left\{\begin{matrix}C\\ø\end{matrix}\right\}C\left\{\begin{matrix}C\\C\ C\\ø\end{matrix}\right\}V$, не равный сегменту вида *VSFSIV*, эквивалентен интервокальному консонантному сегменту, не разделенному слогоразделом, так как это интерлюдия вида $\overline{\left\{\begin{matrix}C\\C\ C\\ø\end{matrix}\right\}\left\{\begin{matrix}C\\ø\end{matrix}\right\}C\left\{\begin{matrix}C\\C\ C\\ø\end{matrix}\right\}}$.

На основе определений (13), (15) и постулата (14) можно заключить, что возможны интервокальные сегменты следующих видов.

1. 2. 1. Сегмент *C* может иметь любую из следующих структур:

(1) $(C = f) \equiv -C$,
(2) $(C = i) \equiv -C$,
(3) $[C = \sim (fVi)] \equiv \overline{C}$,
(4) $[C = (f \cdot i)] \equiv -C$.

1. 2. 2. Сегмент *CC* может иметь любую из следующих структур:

(1) $(CC = ff) \equiv C-C$,
(2) $(CC = fi) \equiv C-C$,
(3) $(CC = ii) \equiv -CC$,
(4) $[CC = \sim (ff \vee fi \vee ii)] \equiv \overline{CC}$,
(5) $[CC = (fi \cdot ff \cdot \sim ii)] \equiv C-C$,
(6) $[CC = (ii \cdot fi \cdot \sim ff)] \equiv (-CC \vee C-C)$,
(7) $[CC = (ii \cdot ff \cdot \sim fi)] \equiv (-CC \vee C-C)$,
(8) $[CC = (ii \cdot ff \cdot fi)] \equiv (-CC \vee C-C)$.

1. 2. 3. Сегмент *CCC* может иметь любую из следующих структур:

(1) $(CCC = fff) \equiv CC-C$,
(2) $(CCC = ffi) \equiv CC-C$,
(3) $(CCC = fii) \equiv C-CC$,
(4) $(CCC = iii) \equiv -CCC$,
(5) $[CCC = \sim (fff \vee ffi \vee fii \vee iii)] \equiv \overline{CCC}$,
(6) $\{CCC = [fff \cdot ffi \cdot \sim (fii \vee iii)]\} \equiv CC-C$,
(7) $\{CCC = [fff \cdot fii \cdot \sim (ffi \vee iii)]\} \equiv (CC-C \vee C-CC)$,
(8) $\{CCC = [fff \cdot iii \cdot \sim (ffi \vee fii)]\} \equiv (CC-C \vee -CCC)$,
(9) $\{CCC = [ffi \cdot fii \cdot \sim (fff \vee iii)]\} \equiv (CC-C \vee C-CC)$,
(10) $\{CCC = [fii \cdot iii \cdot \sim (fff \vee ffi)]\} \equiv (CC-C \vee -CCC)$,
(11) $\{CCC = [fii \cdot iii \cdot \sim (fff \vee ffi)]\} \equiv (C-CC \vee -CCC)$,
(12) $[CCC = (fff \cdot ffi \cdot fii \cdot \sim iii)] \equiv (CC-C \vee C-CC)$,
(13) $[CCC = (fff \cdot ffi \cdot iii \cdot \sim fii)] \equiv (CC-C \vee -CCC)$,
(14) $[CCC \equiv (fff \cdot fii \cdot iii \cdot \sim ffi)] \equiv (CC-C \vee C-CC \vee -CCC)$,
(15) $[CCC = (ffi \cdot fii \cdot iii \cdot \sim fff)] \equiv (CC-C \vee C-CC \vee -CCC)$,
(16) $[CCC = (fff \cdot ffi \cdot fii \cdot iii)] \equiv (CC-C \vee C-CC \vee -CCC)$.

1. 2. 4. Сегмент *CCCC* может иметь любую из следующих структур:

(1) $(CCCC = fffi) \equiv CCC-C$,
(2) $(CCCC = ffii) \equiv CC-CC$,
(3) $(CCCC = fiii) \equiv C-CCC$,
(4) $[CCCC = \sim (fffi \vee ffii \vee fiii)] \equiv \overline{CCCC}$,

(5) $[CCCC = (fffi \cdot ffii \cdot \sim fiii)] \equiv (CCC{-}C \vee CC{-}CC)$,
(6) $[CCCC = (fffi \cdot \sim ffii \cdot fiii)] \equiv (CCC{-}C \vee C{-}CCC)$,
(7) $[CCCC = (fffi \cdot ffii \cdot fiii)] \equiv (CC{-}CC \vee C{-}CCC)$,
(8) $[CCCC = (fffi \cdot ffii \cdot fiii)] \equiv (CCC{-}C \vee CC{-}CC \vee C{-}CCC)$.

Примечание. Структуры типа *CCCC* или *CCC* невозможны в нашей системе, так как они бы противоречили гипотезе *a* (1) о том, что не существует начальных сегментов, содержащих более трех согласных, а *CCCC* противоречила бы также постулату (8) (1. 2).

1. 2. 5. Сегмент *CCCCC* может иметь любую из следующих структур:
(1) $(CCCCC = fffii) \equiv CCC{-}CC$,
(2) $(CCCCC = ffiii) \equiv CC{-}CCC$,
(3) $[CCCCC = \sim (fffii \vee ffiii) \equiv \overline{CCCCC}$,
(4) $[CCCCC = (fffii \vee ffiii)] \equiv (CCC{-}CC \vee CC{-}CCC)$.

Примечание. Структуры — *CCCCC*, *C—CCCC*, *CCCC—C* противоречат гипотезе *a* (1) и поэтому невозможны: структура *CCCCC* невозможна, так как она противоречила бы постулату (8) (1.2).

1. 2. 6. Сегмент *CCCCCC* может иметь лишь следующие две структуры:
(1) $(CCCCCC = fffiii) \equiv CCC{-}CCC$,
(2) $(CCCCCC = \sim fffiii) \equiv \overline{CCCCCC}$.

2. Представим структуру консонантных сочетаний румынского языка на основе системы, изложенной в 1—1. 2. 6.

Каждому из возможных в нашей системе классов или структур сопоставим классы или структуры, существующие в румынском языке. Если какие-либо классы или структуры нашей системы не реализуются в румынском языке, будем обозначать это через ø.

2. 1. В румынском языке существует 20 согласных фонем: *p*, *t*, *k*, *b*, *d*, *f*, *v*, *s*, *z*, *š*, *ž*, *h*, *c*, *č*, *ğ*, *l*, *r*, *m*, *n*. Эти фонемы распределяются в дистрибутивные классы следующим образом:

I_1: *l*, *r*, *n*.
I_2: *p*, *t*, *k*, *b*, *d*, *g*, *f*, *v*, *ž*, *h*, *m*.
I_3: *s*, *z*, *š*
I_4: *c*, *č*, *ğ*.

Примечание: Согласный *č* может употребляться рядом с предначальным согласным *s* или *š* (ср. *'sčenă* и ее вариант *'ščenă*), но никогда не употребляется рядом с собственно начальным согласным.

Согласный *h* не может следовать за предконечным согласным.

F_1: *r*, *l*.
F_2: *p*, *t*, *k*, *b*, *d*, *g*, *f*, *v*, *s*, *z*, *š*, *ž*, *h*, *n*, *m*, *c*, *č*, *ğ*.
F_3: ø

2. 1. 1. Структура начальных сегментов.

2. 1. 1. 1. Сочетания трех согласных могут быть катализованными или некатализованными:

катализованные (1. 1. 2. 1, а): $i_3i_2i_1$: *sbr*, *str*, *skr*, *zdr*, *zgl*, *zgr*, *štr*, *sfr*;

некатализованные (1. 1. 2. 1.):
(1а): $i_{1''}\ i_{1'}\ i_1 : ø$;
(2а): $i_{2''}\ i_2\ i_{2'} : ø$;
(3а): $i_3\ i_{3'}\ i_3 : ø$;

некатализованные (1. 1. 2. 4):

(1а) $i_{2''}i_2i_1 : ø$;
(2а) $i_3i_{3'}i_1 : ø$;
(3а) $i_3i_{3'}i_{2'} : ø$;
(4а) $i_3i_2i_{2'} : ø$;
(5а) $i_3i_{1'}i_1 : ø$;
(6а) $i_{2''}i_{3'}i_{3''} : ø$;

(7a) $i_3 i_2 i_{3''} : ø$;
(8a) $i_{1''} i_2 i_1 : ø$;
(9a) $i_{1''} i_{3'} i_1 : ø$;
(10a) $i_{2''} i_{3'} i_{2'} : ø$;
(11a) $i_3 i_{1'} i_{3''} : ø$.

2.1.1.2. Сочетания двух согласных — это всегда некатализованные сегменты.

(1.1.2.2.2.):

(1a) $i_{1'}$ $i_1 : ø$;
(2a) i_2 $i_{2'}$: *kt*, *kv*, *gv*, *žb*, *žd*, *žg*, *žm*, *žv*;
(3a) i_3 $i_{3'} : ø$.

(1.1.2.2.3.):

(1a) i_2 i_1 : *bl*, *br*, *kl*, *kn*, *kr*, *dr*, *fl*, *fr*, *gl*, *gn*, *gr*, *hl*, *hr*, *žl*, *žn*, *ml*, *mr*, *pl*, *pn*, *pr*, *tr*, *vl*, *vr*;
(2a) i_3 i_2: *sb*, *sk*, *sf*, *sg*, *sm*, *sp*, *st*, *sv*, *šk*, *šf*, *šm*, *šp*, *št*, *šv*, *zb*, *zg*, *zm*, *zv*;
(3a) $i_{3'}$ i_1: *sč*, *sl*, *sn*, *šl*, *šn*, *šr*, *šč*, *zn*;
(4a) i_2 $i_{3''}$: *ps*.

2.1.1.3. Сегменты, состоящие из одного согласного, всегда некатализованы (1.1.2.2.5).

(1a) i_1: *l*, *r*, *n*, *č*;
(2a) $i_{2'}$: *p*, *t*, *k*, *b*, *d*, *g*, *f*, *v*, *ž*, *h*, *m*;
(3a) $i_{3''}$: *s*, *z*, *š*:
(4a) i_a: *c*, *ğ*.

2.1.2. Структура конечных сегментов.

2.1.2.1. Сочетания трех согласных всегда некатализованы.

Сравнивая факты румынского языка с построенной нами системой, мы получаем:

катализованные сегменты (1.1.2.1. b): ø;

некатализованные (1.1.2.2.1):

(1b) f_1 $f_{1'}$ $f_{1''}$: ø;
(2b) $f_{2'}$ f_2 $f_{2''}$: *kst*, *mpt*, *nks*, *nkt*, *stm*;
(3b) $f_{3''}$ $f_{3'}$ f_3: ø

(1.1.2.2.4):

(1a) $f_1 f_2 f_{2''} : ø$;
(2a) $f_1 f_{3'} f_3 : ø$;
(3a) $f_2 f_{3'} f_3 : ø$;
(4a) $f_{2'} f_2 f_3 : ø$;
(5a) $f_1 f_{1'} f_3 : ø$;
(6a) $f_{3''} f_{3'} f_{2''} : ø$;
(7a) $f_{3''} f_2 f_3 : ø$;
(8a) $f_1 f_2 f_{1''} : ø$;
(9a) $f_1 f_{3'} f_{1''} : ø$;
(10a) $f_{2'} f_{3'} f_{2''} : ø$;
(11a) $f_{3''} f_{1'} f_3 : ø$

2.1.2.2. Группы из двух согласных всегда некатализованы (1.1.2.2):

(1b) f_1 $f_{1'}$: *rl*;
(2b) $f_{2'}$ f_2: *bd*, *bt*, *bz*, *ks*, *kt*, *ft*, *gd*, *gn*, *žd*, *mb*, *mf*, *mn*, *mp*, *mt*, *mc*, *nč*, *nk*, *nd*, *nf*, *ng*, *nğ*, *nž*, *ns*, *n*, *nt*, *nc*, *nz*, *pč*, *ps*, *pt*, *sk*, *st*, *sm*, *šk*, *tm*, *zg*, *zm*;
(3b) $f_{3'}$ f_3: ø.

(1.1.2.2.3):

(1b) f_1 f_2: *lb*, *lk*, *lc*, *ld*, *lf*, *lğ*, *lg*, *lp*, *lm*, *ls*, *lš*, *lt*, *lc*, *lv*, *lz*, *rb*, *rk*, *rč*, *rd*, *rf*, *rg*, *rğ*, *rh*, *rm*, *rn*, *rp*, *rs*, *rš*, *rt*, *rc*, *rv*, *rz*;
(2b) $f_2 f_3$: ø;
(3b) $f_1 f_3$: ø;
(4b) $f_{3''} f_2$: ø.

2.1.2.3. Сегменты, состоящие из одного согласного, всегда некатализованы.

(1b) f_1: *lr*;

(2b) $f_{2'}$: *p, t, k, b, d, g, f, v, s, z, š, ž, h, č, ğ, c, m, n*;
(3b) $f_{3''}$: ø;
(4b) f_{a}: ø.

2.2. Структура интервокальных сегментов[2].

Интервокальные сегменты могут иметь вид *C, CC, CCC. CCCC, CCCCC*. Следовательно, отсутствуют сегменты вида *CCCCCC*, существующие в изложенной выше метасистеме.

2.2.1. Сегмент вида *C* может обладать любой из следующих структур (1.2.1.):

(1) —*C*: *p, t, k, b, d, g, f, v, s, z, š, ž, h, c, č, ğ, l, r, m, n*;
(2) —*C*: *p, t, k, b, d, g, f, v, s, z, š, ž, h, c, č, ğ, l, r, m, n*;
(3) —$\overline{C}$: ø;
(4) —*C*: *p, t, k, b, d, g, f, v, s, z, š, ž, h, c, č, ğ, l, r, m, n*.

2.2.2. Сегмент вида *CC* может обладать любой из следующих структур (1.2.2):

(1) *C* — *C*: ø;
(2) *C*—*C*: *bz, bč, bž, bp, bs, bš, bt, bv, kb, kč, kd, kğ, km, kp, kc, kz, čn, dk, dg, dm, dn, ds, dv, fğ, fš, gm, gz, hn, lg, lh, ll, ln, mf, mb, mm, mv, mz, nh, nl, nm, nn, np, nr, nv, rž, sž, sš, sc, sz, tb, tk, td, ck, vh, vš, vz, zr*;
(3) *CC*: ø
(4) $\overline{CC}$: ø
(5) *C*—*C*: *bt, ks, gd, lb, lč, ld, lf, lg, lp, lš, lc, lv, lz, mb, mf, mp, nk, nč, nd, nf, ng, nğ, nž, ns, nš, nt, nc, pč, rb, rk, rč, rd, rf, rg, rh, rl, rm, rn, rp, rs, rš, rt, rc, rv, rz, tm, tč*;
(6) (—*CC* · *C*—*C*): *bl, br, kl, kn, kr, kv, dr, fl, fr, gl, gr, hl, hr, žb, žg, žl, žm, žn, pn, pr, sf, sl, sn, sp, šč, tr, vr, zb, zv*;
(7) (—*C* · *C*—*C*) = ø;
(8) (—*CC* · *C*—*C*): *žd, ps, sk, sm, st, šk, zm*.

2.2.3. Сегмент *CCC* может иметь любую из следующих структур (1.2.3.):

(1) *CC*—*C*: ø
(2) *CC*—*C*: *ksh, ldm, lds, lpt, ltk, ltč, ltf, ltm, mp, nkm, nčs, nčz, ndv, ngš, nsh, nss, ntd, ntl, ptn, pts, ptz, rbk, rks, rkš, rdm, rdn, rfš, rgš, rtk, rtf, rth, rtž, rtn, rtp, rts, rtv, stf, stn, stp, stv*;
(3) *C*—*CC*: *bsk, bsč, bst, bšt, zvr, bgr, bpr, fdr, msp, nml, npr, nsd, nvr, tgr, tšt, zbl*;
(4) —*CCC*: ø;
(5) $\overline{CCC}$: ø;
(6) *C*—*CC* · *CC*—*C*: *mpt, nks, stm, rtm*;
(7) *CC*—*C* · *C*—*CC*: ø;
(8) *CCC*—*C* · —*CCC*: ø;
(9) *CC*—*C* · *C*—*CC*: *ktr, lkl, ltr, mbl, mbr, mff, mpl, ndr, ngl, ngr, nst, nšt, ntr, ptr, rpl, rpr, rsp, rst, rsl, rtr, ksk, ksč, ksm, ksp, lpn, lsp, nkl, nkr, nfl, nfr, ngv, nžg, nsk, nsč, nsf, nsl, nsm, nsp, nsv, nšf, psk, rkt, rgr,ršr*;
(10) *CC*—*C* · —*CCC*: ø:
(11) *C*—*CC* · —*CCC*: *štr, zgl, zgr, spr*;

---

[2] Ниже используется материал, содержащийся в исследовании Александры Рочерик-Александреску (A. Roceric-Alexandrescu. Cîteva observaţii asupra semnalelor demorcative în limba romînă. «Fonetică şi dialectologie», vol. IV — в печати). Указанное исследование содержит перечень интервокальных сочетаний согласных, установленный на материале словаря «Dicţionarul limbii romîne literare contemporanc» (Bucureşti, 1954), а также на материале работы: A. Graur, A. Rosetti. Esquisse d'une phonologie du roumain. «Bulletin linguistique», VI, p. 5 и словаря «Dicţionar invers» (Bucureşti, 1955).

(12) *CC—C · C—CC*: *kst*, *nkt*;
(13) *CC—C · —CCC*: ø;
(14) *CC—C · C—CC*: ø;
(15) *CC—C · C—CC · —CCC*: *skl*, *skr*, *spl*, *spr*, *str*, *zdr*;
(16) *CC—C · C—CC · —CCC*: ø.

2. 2. 4. Сегмент вида *CCCC* может обладать любой из следующих структур (1. 2. 4):

1. *CCC—C*: ø;
2. *CC—CC*: *rkgr*, *rtdr*
3. *C—CCC*: *lstr*, *mskr*;
5) *CCC—C · CC—CC*: ø;
6) *CCC—C · C—CCC*: ø;
7) *CC—CC · C—CCC*: *kskb*, *kskr*, *kspl*, *kspr*, *kstr*, *nskr*, *nspr*, *nstr*;

(4) $\overline{\text{CCCC}}$: *bštn*, *rstn*;
(8) *CCC—C · CC—CC · C—CCC*: ø.

2. 2. 5. Сегмент вида *CCCCC* может обладать любой из следующих структур (1. 2. 5.):

(1) *CCC—CC*: ø;
(2) *CC—CCC*: *nčsrt*, *ptspr*;
(3) $\overline{\text{CCCCC}}$: ø;
(4) *CCC—CC · CC—CCC*: ø.

2. 2. 6. Сегменты вида *CCCCCC* не существуют в румынском языке, что выражается формулами (1. 2. 6.):

(1) *CCC—CCC* = ø,
(2) $\overline{\text{CCCCCC}}$ = ø.

THE STRUCTURE OF CONSONANT GROOPS IN RUMANIAN

Summary

The first part of the paper deals with a series of theoretic principles. It formulates a system of formalised definitions given to the consonantical classes of distribution: initials, postinitials, preinitials; finals, prefinals, postfinals. Then the secundary and tertiary functions of the consonants are defined.

At last, the place where the syllable cut may occur is established takind account of the structure of the intervocalic clusters.

The second part of the paper applies the theory formulated in the first part to the consonant system of Rumanian.

III. ГРАММАТИКА И ЛЕКСИКОЛОГИЯ

И. И. Ревзин

ОПЫТ СИМВОЛИЧЕСКОГО ИЗЛОЖЕНИЯ НЕМЕЦКОГО СИНТАКСИСА ПРИМЕНИТЕЛЬНО К ПРОБЛЕМАМ МАШИННОГО ПЕРЕВОДА

§ 1. Поскольку электронно-вычислительные машины, на которых должен производиться машинный перевод, действуют на основе программ, составленных по принципам символической (или математической) логики, представляется целесообразным попытаться изложить основные правила синтаксиса символическим образом, что позволило бы:

а) упростить по возможности процесс программирования (особенно в части исключения грамматической многозначности и омонимии);

б) установить формальные критерии для определения того, верно или неверно анализируется предложение в машине;

в) проверить формальные границы предложения (т. е. тех минимальных отрывков текста, в пределах которых производится анализ — так называемая «проблема законченности»).

При создании соответствующих символизированных правил мы будем идти следующим путем: а) постараемся на основе общих приемов, применяемых в символической логике, построить некую исходную абстрактно-логическую систему, которой вполне могла бы оперировать машина, и б) путем введения соответствующих обозначений и формально-определенных операций преобразования попробуем привести к нашей исходной системе законы немецкого синтаксиса.

А. СОЗДАНИЕ ИСХОДНОЙ СИСТЕМЫ

§ 2. Мы будем рассматривать некоторые элементы $E(E_1, E_2, \ldots, E_n)$. Все элементы мы делим на два класса S и V. Далее снабжаем каждый элемент класса S и класса V индексами множественности *en* (*pl* или *sing*), которые подписываются снизу, а элементы класса S еще и индексами отношения (N, A и некоторыми другими), которые надписываются сверху. Позднее вводятся и некоторые другие индексы, например индексы направления (стрелки), указывающие, что элемент вступает во взаимодействие с элементом, стоящим справа или слева.

Отдельные элементы могут объединяться. Объединение элементов (K) получается путем конъюнкции и дизъюнкции (правда, последняя — в несколько видоизмененном значении по сравнению с тем, которое придается этому виду связи в символической логике[1]). Однако, по-

[1] См., например: Д. Гильберт и А. Аккерман. Основы теоретической логики. М., 1947, гл. I.

скольку в дальнейшем строится определенная система реляционных значений, определение истинности или ложности соответствующего элемента дается несколько отличное. Мы будем считать элемент истинным, если он находится в отношении подчинения к заданному элементу E или к заданной конструкции K. Поскольку это определение лежит в основе всей системы, нужно сразу же подчеркнуть одно важное обстоятельство. Хотя вначале, при введении основных соотношений, мы будем считать истинность или ложность элемента заранее установленной, следует помнить, что в конкретном применении этих соотношений элемент истинен или ложен не вообще, а только по отношению к заданному элементу или к заданной конструкции. Установление истинности происходит на основе правил, изложенных в § 5.

Значение функции истинности мы будем считать равным единице (символически это обозначается $f(E_i)=1$) для истинного элемента и равной нулю ($f(E_i)=0$) для ложного элемента.

Переходим к введению основных соотношений.

§ 3. Под конъюнкцией (мы будем обозначать ее знаком умножения) в символической логике понимается такое объединение элементов, которое истинно в том и только том случае, если истинны все его элементы.

$E_k \times E_l$ истинно, если истинно и E_k и E_l.

Под дизъюнкцией (мы будем обозначать ее знаком сложения) в символической логике понимается такое объединение элементов, которое истинно в том случае, если истинен хотя бы один из его элементов. Мы введем дополнительно одно ограничение, а именно—истинный элемент в дизъюнкции должен быть единственным, а для этого договоримся, что 1) как только нам на основании каких-то сведений стало известно, что один из членов дизъюнкции оказался истинным, то все остальные члены дизъюнкции мы считаем ложными и отбрасываем; 2) в любой дизъюнкции все члены должны быть разные.

Если функция истинности данного элемента установлена, его можно заменять значением (единицей или нулем). Имеют место следующие соотношения математической логики:

$$E \times E = E, \tag{1}$$
$$E \times 1 = E, \tag{2}$$
$$E \times 0 = 0, \tag{3}$$
$$E + 0 = E. \tag{4}$$

Эти соотношения следуют из определения конъюнкции и дизъюнкции. Несколько отличаются от обычных соотношений математической логики следующие, введенные для обеспечения ограничения на дизъюнкцию:

$$E_1 + E_2 = \begin{cases} E_2, & \text{если } f(E_2)=1, \\ E_1, & \text{если } f(E_2)=0. \end{cases} \tag{5}$$

$$[E+1=1] \to [f(E)=0]. \tag{6}$$

Устанавливаем, что

$$E_i \times (E_k + E_l) = (E_i \times E_k) + (E_i \times E_l). \tag{7}$$

Это дистрибутивный закон, аналогичный такому же соотношению в алгебре.

Для дизъюнкции действует коммутативный закон:

$$E_i + E_k = E_k + E_l. \tag{8}$$

Для конъюнкции в рассматриваемой системе коммутативный закон, вообще говоря, не действует. Это связано с тем, что при перестановке в конъюнкции истинность одного из элементов по отношению к какому-то третьему элементу может измениться. Поэтому мы каждый раз будем оговаривать некоммутативность выражения.

Для конъюнкции действует ассоциативный закон:

$$E_i \times (E_k \times E_l) = (E_i \times E_k) \times E_l = E_i \times E_k \times E_l. \qquad (8')$$

Если установлена истинность или ложность составляющих элементов, то в объединении на основании соотношений 4, 5, 6, 8 и 8′ могут быть удалены все дизъюнктивные члены, за исключением одного. Останется, таким образом, только конъюнкция. Выражение, приведенное к такой форме, мы назовем нормальной, или конъюнктивной, формой.

§ 4. В нашей системе мы будем рассматривать также некоторые элементы A, именуемые атрибутами. Атрибут — это свойство, характеризующее один из какого-то ряда элементов E_i и могущее в конкретном объединении быть идентифицировано с ним. Поэтому мы определяем атрибут соотношением:

$$A = E_1 + E_2 + \ldots + E_m, \qquad (9)$$

причем все элементы правой части имеют различные индексы.

Относительно атрибутов мы можем получить следующее, может быть, тривиальное, но важное для дальнейшего предложение:

Конъюнкция атрибута с одним из элементов, характеризуемых им, равна этому элементу.

Нужно доказать, что $A_i \times E_i = E_i$. (10)

Это равенство равносильно, согласно (9), следующему:

$$(E_1 + E_2 \ldots + E_i + \ldots + E_m) \times E_i = E_i.$$

Если E_i истинно, то выражение

$$(E_1 + E_2 + \ldots + E_i + \ldots + E_m)$$

содержит истинный член E_i, значит, остальные могут быть отброшены. Имеем $E_i \times E_i = E_i$, но это соотношение (1). Предположим теперь, что E_i ложно, но тогда в левой части его ложный конъюнктивный член, и конъюнкция всегда ложна, вне зависимости от того, ложно или истинно A_i.

Эквивалентность (10) доказана. Ее легко можно распространить на случаи типа $A_i \times A_i \times E_i = E_i$ и т. д.

Атрибуты, характеризующие V, мы будем обозначать A^V, а атрибуты, характеризующие S, — A^S. Имеем, согласно (10):

$$A^V \times V = V, \qquad (11)$$

$A^S \times S = S$ (второе, вообще говоря не коммутативно, так как зависит от индекса направления у A^S).

Для атрибутов второго ранга (т. е. атрибутов атрибутов) A^{A^S} и A^{A^V} имеем:

$A^{A^S} \times A^S = A^S$ (первое никогда не коммутативно),

$A^{A^V} \times A^V = A^V$ (11′)

§ 5. Теперь мы можем перейти к формулировке правил определения истинности элементов.

Первое правило. Если атрибут первого ранга, при заданном S или V, имеет индекс направления, указывающий в сторону этого S

или V, то он вступает с ним в соотношения (10), и мы будем считать этот атрибут истинным по отношению к заданному элементу S или V. Аналогично атрибут второго ранга, вступающий в соотношения (11′) с A^V или A^S, мы будем считать истинным по отношению к заданному элементу A^V или A^S.

Второе правило. Если мы имеем объединения типа:

$$S^N_{en} \times V_{en} \text{ или } S^N_{en} \times N_{en} \times S^A_{en} \text{ или } V_{en} \times S^N \tag{12}$$

или можем прийти к ним в результате определенных преобразований, и если у V и S^N одинаковые индексы множественности (см. § 2), то оба (или все три) элемента считаются истинными по отношению к соответствующему заданному объединению. Само объединение мы будем считать правильным (или высказыванием) и будем обозначать его в дальнейшем формулой: Connex (S, V) или Connex $(S, V) \times S^A$. Функция высказывания принимает в этом случае значение 1, в противном случае — значение 0:

$$\mp(\text{Connex } (S, V)) = 1.$$

У нас будут в дальнейшем встречаться дизъюнкции элементов типа $S^N_{en} + S^A_{en}$. Для удобства мы будем иногда рассматривать такое объединение как единый элемент, написанный в форме S^{N+A}_{en}, и говорить, в целях краткости, что индекс содержит дизъюнкцию. Кроме того, относительно всех индексов условимся: если сверху или снизу какой-то индекс не пишется, то это означает, что на соответствующем месте может стоять любой индекс, т. е. что для данного соотношения соответствующий индекс не важен. Введем еще следующие правила («аксиомы порядка»):

$$S^{N+A}_{en} \times V_{en} \times S^{N+A} = S^N_{en} \times V_{en} \times S^A, \tag{13}$$

$$S^{N+A} \times V_{en} = S^N_{en} \times V_{en}. \tag{13′}$$

§ 6. Устанавливаем, что соответствующие элементы или имеют индексы с самого начала, или получают их в результате операции переноса ($\ddot{u}$, т. е. Übertragung). Оператор переноса ($\ddot{u}$), например, в виде: $\ddot{u}\left(\frac{N}{E_i} + \frac{N}{E_2} \ldots + \frac{A}{E_n}\right)$, будучи применен к конюнктивному члену E_i, стоящему вправо от него, переносит на него соответствующий индекс (этот индекс записан в числителе, а элемент, на который он может быть перенесен — в знаменателе), таким образом:

$$\vec{\ddot{u}}\left(\frac{N}{E_1} + \frac{A}{E_2} + \frac{N}{E_3} + \ldots + \frac{A}{E_1} + \ldots \frac{N}{E_n}\right) \times E_1 = E^N_1. \tag{14}$$

Левая часть формулы (14) не коммутативна.

Кроме того, существует условный оператор переноса ($b\ddot{u}$ — bedingte Übertragung), который применяется к элементу E_i так же, как и оператор ($\ddot{u}$), но в том и только том случае, если влево от него не стоял оператор ($\ddot{u}$). Таким образом, оператор ($b\ddot{u}$) действует в выражениях типа

$$\vec{b\ddot{u}}\left(\ldots + \frac{N}{E_i} + \ldots\right) \times E_i = E^N_i$$

и не действует в выражениях типа

$$\vec{\ddot{u}}\left(\ldots + \frac{N}{E_i} + \ldots\right) \times \vec{b\ddot{u}}\left(\ldots + \frac{A}{E_i} + \ldots\right) \times E_i = E^N_i$$

(а не E^A_i или E^{N+A}_i).

§ 7. Кроме того, вводится ряд операторов превращения.

1. Оператор pr, который встречается в выражениях (или частях выражений):

$$S \times \overrightarrow{pr} \times S^{D+A}$$

или

$$V \times \overrightarrow{pr} \times S^{D+A}$$

или

$\overrightarrow{pr} \times S^{D+A} \times V$ (такие выражения не коммутативны) и, будучи применен к S^{D+A}, превращает его в первом выражении в элемент природы A^S:

$$S \times \overrightarrow{pr} \times S^{D+A} = S \times \overleftarrow{A^S}, \tag{15}$$

а во втором и третьем выражении — в элемент природы A^V:

$$V \times \overrightarrow{pr} \times S^{D+A} = V \times A^V. \tag{15'}$$

Если оператор pr применен к заданному элементу S^{D+A}, мы будем считать pr истинным по отношению к S^{D+A}, а последний, в свою очередь, истинным по отношению к S в (15) и соответственно V в (15').

2. Операторы Q_S и Q_A, которые применяются к некоторым выражениям Connex (а именно — только выражениям типа $S^N_{en} \times V_{en}$, а также к не имеющим в других случаях смысла выражениям $S^N_{en} \times S^A \times V_{en}$), и превращают их соответственно в элементы природы A^S или A^V. Таким образом:

$$\overrightarrow{Q_S} \times \mathrm{Connex}(S,\ V) = A^S, \tag{16}$$

$$\overrightarrow{Q_S} \times \mathrm{Connex}(S,\ V) \times S^A = A^S, \tag{16'}$$

$$\overrightarrow{Q^V} \times \mathrm{Connex}(S,\ V) = A^V, \tag{17}$$

$$\overrightarrow{Q_V} \times \mathrm{Connex}(S,\ V) \times S^A = A^V. \tag{17'}$$

Если оператор Q применен к заданному Connex (S, V), мы будем считать Q истинным по отношению к этому объединению.

3. Операторы I' и I'', G' и G'' построены несколько отлично от других операторов. Элемент под знаком такого оператора рассматривается как V (без индекса множественности), но после того, как согласно (10) с этим V вступили в взаимодействие все относящиеся к нему атрибуты, оператор I или G превращает этот элемент в элемент природы A^V или A^S (подробнее см. § 11, пункт 13). Условно эта особенность оператора отражается символом &, например $I'(V \,\&\, A^V)$ — подчеркиваем, что & не знак связи типа $+$ или $\times$, а просто условное обозначение.

Если A^V или A^S, получившиеся в результате действия оператора, в дальнейшем вступает в соотношения (10) в качестве атрибута, мы будем считать элемент под знаком оператора истинным по отношению к тому элементу, с которым A^V или A^S вступает в соотношения (10).

Введем теперь последовательность применения операторов превращения.

1. Оператор pr применяется первым, если непосредственно за элементом, к которому он применяется, не следует оператор Q_S, в противном случае первым применяется оператор Q_S;

2. Преобразование элемента под знаком оператора I'' в A^V или A^S производится первым, если этот оператор следует за элементом с индексом (i), что же касается элемента V, то он с самого начала выведен из-под знака оператора, эти правила его не касаются.

3. Если оператор Q стоит между элементом с индексом (i) и оператором I'', он применяется первым.

4. Если оператор I'' стоит за оператором Q, а перед Q нет элемента с индексом (i), то I'' применяется первым.

Если, наконец, все эти условия не соблюдаются, операторы применяются сначала в порядке следования в выражении, а если высказывания не получилось, пробуются другие комбинации, пока в конце концов окончательное выражение не становится высказыванием.

Условимся еще, что элемент, стоящий непосредственно перед V_{en}, теряет индекс направления влево.

Последнее замечание — чисто технического свойства: в целях экономии записи мы никогда не будем ставить скобок в выражениях типа $E_i \times E_k + E_l$, которые будут означать $(E_i \times E_k) + E_l$, иначе: мы уславливаемся, что конъюнкция связывает теснее, чем дизъюнкция.

Необходимая нам для дальнейшего исходная система построена. Она, по-видимому, легко может быть реализована в машине.

Б. ЛИНГВИСТИЧЕСКОЕ ИСТОЛКОВАНИЕ ПОСТРОЕННОЙ СИСТЕМЫ

§ 8. Пора объяснить лингвистический смысл построенной системы. Под ее элементами следует понимать грамматические значения соответствующих форм слова, например S^N_{pl} — значение существительного именительного падежа множественного числа, V — значение глагола, A^S — значение прилагательного[2], A^{A^S} или A^V — наречия, pr — предлога. Однако не всегда истолкование символа соответствует общепринятой терминологии. Дело в том, что соответствующие категории должны выделяться строго по структурным принципам, поэтому, например, наречие может получить два разных символа.

Оператор переноса появляется там, где грамматическое значение одного слова зависит от других слов, которые сами по себе не несут значения (артикли) или одновременно сами имеют значение (прилагательные, которые значением своей склоняемой формы определяют значение следующего за ним существительного — таков, например, смысл нашего условного оператора переноса). Оператор pr, как уже указано, не что иное, как предлог, при помощи этого оператора предложная группа может быть связана с глаголом или существительным. Операторы Q связывают придаточное предложение со сказуемым главного предложения или существительным, к которому оно относится. Наконец, операторы I и G отражают природу инфинитивов и причастий (ср. § 11, пункт 13).

«Аксиомы порядка» (13 и 13′) отражают особенности порядка слов и позволяют установить грамматическое значение в случаях типа *Мать любит дочь* и *Дочь любит мать* (как известно, в немецком синтаксисе такого рода формальное совпадение подлежащего и дополнения — случай весьма распространенный). При определении порядка элементов в высказывании, к которому применен оператор Q, отражена особенность порядка слова в придаточном предложении немецкого языка. Вообще все соотношения, не взятые непосредственно из логики, приспособлены к тем или иным синтаксическим явлениям.

§ 9. Следует особо остановиться на лингвистическом смысле введенных понятий конъюнкции и дизъюнкции[3]. Как известно, одна форма

---

[2] Конечно, может вызвать сомнение правомерность трактовки прилагательного как атрибута в том смысле, который указан в § 4. Я исходил из того соображения, что прилагательное характеризует свойство, присущее ряду существительных, но в конкретной ситуации относится к одному и даже может заменить его, например: *Там лежат разные карандаши. Дай мне зеленый.*

[3] Здесь мы следуем той лингвистической интерпретации конъюнкции и дизъюнкции, которая предложена Л. Ельмслевом в «Пролегоменах к теории языка» («Новое в лингвистике», I, М., 1960, стр. 295—297).

слова может иметь целый ряд значений, лишь одно из которых реализуется в соответствующих константах, так, например, *lange* может иметь значение атрибута к существительному (прилагательное) и атрибута к глаголу (наречие), *das* может иметь два значения: существительного (или по обычной терминологии — местоимения), значение союзного слова и т. п.

Смысл дизъюнкции и состоит в нашем случае в объединении всех значений данной формы слова. Конъюнкция употребляется в том случае, когда одна форма одновременно передает два значения, например, *das* в значении союзного слова: *das* здесь и показатель подчинения (оператор Q_s) и подлежащее или дополнение. В нашей символике *das* изобразится поэтому как

$$das = \ddot{u}\left(\frac{N+A}{S_2}\right) + S_2^{N+A} + \left[Q^S(\ ) \times S_{sing}^{N+A}\right]$$

значение артикля — *местоимение* — *союзное слово*

Основная роль конъюнкции в нашей системе отводится, однако, другому: при ее помощи объединяются совокупности значений всех слов данного предложения. Это вполне соответствует определению конъюнкции как связи, истинной только в том случае, если каждый ее член вступает в определенные отношения с другими.

Смысл конъюнктивной или нормальной формы состоит в том, что в результате некоторых операций исчезает (выраженная количеством дизъюнкций [4]) неопределенность в определении значений данной формы слова и все элементы высказывания становятся однозначно определенными (речь идет, конечно, о грамматических значениях).

Таким образом, смысл всех операций состоит в нижеследующем. Мы полагаем все объединение высказыванием, затем проделываем все операции и устанавливаем истинность или ложность всех элементов. Ложные элементы дизъюнкции (т. е. реляционные значения, которых слово в данном предложении не имеет) таким путем удается исключить. Если же ложным в нормальной форме оказался хотя бы один член конъюнкции, то наша посылка о правильности выбранного нами предложения не верна или же неправильно выбран порядок проведения операций. Практически это означает, что нужно вернуться к определенному этапу и произвести действия в другом порядке или же что границы предложения определены неверно и необходимо передвинуть его конец вправо или влево (здесь важно отметить наличие дополнительных указаний на границу предложения в виде знаков препинания).

§ 10. Как мы уже указали, элементами нашей системы являются не слова, как они выступают в словаре, а грамматические значения форм слов, как они берутся из разных связанных текстов.

Поскольку мы построили чисто реляционную систему, то мы не вправе касаться вопросов, как формируются соответствующие элементы. Предполагается, что машина нашла в словаре основы слов переводимого текста, выделила окончания или другие форманты и по взаимодействию основы слова с формальным элементом определила (например, при помощи соответствующей таблицы) возможные грамматические значения данной формы слова.

Нет в построенной системе места и для так называемых аналити-

[4] Точнее: неопределенность зависит от количества дизъюнкций и вероятности для каждого дизъюнктивного члена оказаться истинным в данном высказывании. Однако этот вопрос, связанный с введением меры грамматической информации, нуждается в особом рассмотрении. Ср. замечания Р. Якобсона в работе «Linguistics and communication theory» («Structure of language and its mathematical aspects». «Proceedings of American Mathematical Society», vol. XII, Providence, 1961, стр. 251).

ческих глагольных форм (*hat*, *wird beschlossen* и т. п.). И это понятно. В качестве элементов у нас выступают только синтаксические (вернее, реляционные), а ни в коем случае не чисто морфологические значения, как, например, значение времени или наклонения, и наша система только для того и служит, чтобы помочь выяснить именно реляционные значения. Впрочем, поднятый вопрос об аналитических формах носит чисто теоретический характер. Практически вполне возможно построить соответствующие операторы переноса, которые и решат проблему аналитических форм. В интересах логической строгости мы этого, однако, делать не будем, а предположим, что уже сформулированы какие-то правила, по которым мы можем объединить оба члена аналитической формы и снабдить их единым индексом (V_{sing} или V_{pl}).

В. СИМВОЛИЗАЦИЯ НЕМЕЦКОГО СИНТАКСИСА

§ 11. Теперь мы можем приступить к выполнению нашей основной задачи — формализации правил немецкого синтаксиса. Покажем, прежде всего, основные элементы, соответствующие различным классам слов. При этом в целях большей наглядности мы будем исходить по возможности из традиционного деления на части речи.

Существительные

Для существительных единственного числа вводятся добавочные индексы 1, 2, 3, обозначающие соответственно существительные мужского, среднего и женского рода. Условимся, что, если стоит один из этих индексов, то индекс *sing* не пишется.

1. Формы типа *Tag*, *Bach*, *Wald*, *Haar*:
$Tag = S_1^{N+D+A}$; $Tage = S_1^D + S_{pl}^{N+G+A}$; $Tagen = S_{pl}^D$; $Tages = S_1^G$; $Haar = S_2^{N+D+A}$ и таким же образом далее с заменой индекса 1 на 2.

2. Формы типа *Kuh*:
$$Kuh = S_3^{N+D+A};\ K\ddot{u}he = S_{pl}^{N+A+G};\ K\ddot{u}hen = S_{pl}^D.$$

3. Формы типа *Adler*:
$$Adler = S_1^{N+G+D+A} + S_{pl}^{N+G+A};\ Adlern = S_{pl}^D,\ Adlers = S_1^G.$$

Принцип из этих примеров ясен, так же индексируются и существительные других типов. Отметим только, что слова типа *Möglichkeit*, *Notwendigkeit*, т. е. слова, к которым может относиться инфинитивный оборот, получают еще особый индекс (i).

Местоимения и артикли

4. Формы типа *er*, *sie*:
$$er = S_{sing}^N,\ ihnen = S_{pl}^D.$$

5. Формы типа *mein*, *dieser*:
$$mein = A^S \times \vec{u}\left(\frac{N}{S_1} + \frac{N+A}{S_2}\right);\ diesem = A^S \times \vec{u}\left(\frac{D}{S_1+S_2}\right).$$

6. Формы типа *der*, *die*, *das*:
$$der = \vec{u}\left(\frac{N}{S_1} + \frac{G+D}{S_3} + \frac{G}{S_{pl}}\right) + \vec{Q}_A\left(\frac{A}{S_1+S_3+S_{pl}}\right) \times S_{sing} + S_{sing}^N;$$

второе значение (выраженное двумя элементами, связанными конъюнкцией) — значение союзного слова, третье — значение указательного местоимения.

7. Формы типа *es*:

$$es = {}^{(i)}S^{N+A}_{sing};$$

индекс (i) обозначает, что вправо от *es* может стоять инфинитивный оборот или придаточное предложение, относящееся к нему.

Прилагательные и наречия

8. Формы типа *gut*:

$$gut = A^V.$$

9. Формы типа *lange*:

$$lange = A^V + \left(\vec{A}^S \times b\ddot{u}\left(\frac{N+A}{S_3} + \frac{N+A}{S_{pl}}\right) + A^S\right).$$

Первое значение — наречие (например, в приводимом ниже примере из романа Г. Манна); *bü* — оператор условного переноса, он работает в случаях типа: *Er sah* ***lange*** *Menschenschlangen vor dem Arbeitsamt* и не работает в случаях типа: *die lange Kette*. A^S пишется два раза в силу формальной необходимости, ведь если оператор *bü* не был применен, т. е. ложен, то ложна и вся конъюнкция $A^S \times b\ddot{u}(\ )$. Тогда, однако, останется просто A^S.

10. Формы типа *klaren*:

$$klaren = \vec{A}^S \times b\ddot{u}\left(\frac{A}{S_1} + \frac{D}{S_{pl}}\right) + A^S;$$

четыре формы типа *sehr*:

$$sehr = \vec{A}^{A^S}$$

Глаголы

11. Формы типа *macht*, *liebte*, *könnte*, *ist gekommen* (ср. замечание об аналитических формах):

$$macht = V_{sing}.$$

12. Формы типа *liebten*, *waren*, *kommen*:

$$liebten = V_{pl}.$$

13. Формы типа *lieben*:

$$lieben = V_{pl} + I'(V \,\&\, A^V);$$

Второе значение — значение инфинитива — передается оператором I'. Смысл его состоит в том, что в предложениях типа *er kann gut arbeiten* сначала объединяются *gut* с *arbeiten* как A^V с V, затем полученный элемент превращается в A^V). Это соответствует двойственной природе инфинитива как формы, к которой относятся некоторые формы и которая сама относится к другому слову, или, иначе, инфинитива как двойственной глагольно-именной категории.

14. Формы типа *lieben*, *aufzustehen*:

$$aufzustehen = I''(V \,\&\, A^V + A^S).$$

Оператор I'' аналогичен оператору I'; введение двух различных операторов позволяет отразить некоторые отличия между обеими формами (например, то, что инфинитив с *zu* может быть атрибутом к существительному), на которых мы не имеем возможности останавливаться.

15. Формы типа *verlockend*:

$$verlockend = G'(V \,\&\, A^S).$$

Оператор G' вполне аналогичен оператору I', т. е. с реляционной точки зрения причастия функционируют так же, как и инфинитивы.

16. Формы типа *geleistet* (исключая случай, когда причастие II входит в состав аналитической формы):

$$geleistet = G''(V \,\&\, A^S).$$

И этот оператор аналогичен предыдущим. Введение различных символов оставляет возможность для введения правил образования аналитических форм.

17. Формы типа *geleistete*, *verlockendem*:

$$geleistete = b\ddot{u}\left(\frac{N+A}{S_3} + \frac{N+A}{S_{pl}}\right) \times G''(V \,\&\, A^S); \quad verlockendem = \ddot{u}\left(\frac{D}{S_1 + S_2}\right) \times G''(V \,\&\, A^S).$$

Предлоги

18. Формы типа *in*:

$$in = \vec{pr}.$$

19. Формы типа *vor*, *nach*:

$$vor = \vec{pr} + \bar{T}^V.$$

Второе значение — значение отделяемой приставки. С формальной точки зрения этот элемент действует как наречие A^V. Однако, поскольку в словаре берется перевод всего глагола, целесообразно ввести специальное обозначение. Впрочем, оно имеет и некоторый формальный смысл: объединение T^V с V_{en} должно производиться раньше всех других операций $A^V \times V$, причем T^V объединяется только с V_{en}, стоящим влево от него.

20. Форма *zu*:

$$zu = \vec{pr} + \bar{T}^V + \vec{u}\left(\frac{I''}{V}\right).$$

Последнее значение и есть значение оператора I'', рассмотренного выше; глагол получает его в случае, когда *zu* не стоит внутри него.

Союзы

21. Формы типа *und*, *aber*:

$$und = К$$

К — своеобразный условный оператор, говорящий о том, что последующий элемент надо объединить с предыдущим, имеющим те же индексы (не учитываются лишь индексы множественности), так как союз объединяет однородные члены предложения. Если же это невозможно, К — граница предложения, т. е. союз объединяет два предложения. Без этого оператора, как кажется, можно обойтись, отбросив при анализе К и применяя соотношение (1). Поэтому мы не рассматривали его в разделе А.

22. Формы типа *weil*, *daß*:

$$weil = \bar{Q}^V(\text{Connex } (V,\ S)).$$

Оператор устанавливает, что если S и V связаны, всю группу надо превратить в A^V, т. е. что придаточное относится к сказуемому главного предложения.

$$da\beta = \bar{Q}^{V+S}(\text{Connex } (V,\ S)) —$$

то же самое, но учитывается, что придаточное с *daß* может относиться к существительному.

§ 12. Покажем теперь на одном предложении, взятом из романа Г. Манна «Верноподанный», как действует построенная система. Для наглядности элементы будут подписываться непосредственно под словами:

Sie liebte es, vor einem Bild,

$$(S_{pl}^{N+A} + S_{sing}^{N+A}) \times V_{sing} \times {}^{(i)}S_{sing}^{N+A} \times (\vec{pr} + \vec{T}^V) \times \vec{u}\left(\frac{D}{S_1 + S_2}\right) \times S_2 \times$$

das ihr

$$\times\left[\vec{u}\left(\frac{N+A}{S_2}\right) + \vec{Q}\left(\frac{A}{S_2}\right) \times S_{sing}^{N+A}\right] \times \left[A^V + \vec{A}^S \times \vec{u}\left(\frac{N}{S_1} + \frac{N+A}{S_2}\right)\right] \times$$

gefiel, lange

$$\times V_{sing} \times \left[A^V + \left(\vec{A}^S \times b\vec{u}\left(\frac{N+A}{S} + \frac{N+A}{S} + A^S\right)\right]\right] \times$$

stehenzubleiben mit halbgeschlossenen Augen

$$\times I''(V \,\&\, A^V + \vec{A}^S) \times (\vec{pr} + \vec{T}^V) \times \left[A^S \times b\vec{u}\left(\frac{A^S}{S_1} + \frac{D}{S_{pl}}\right) + \vec{A}^S\right] \times S_{pl}$$

В этом выражении есть конъюнктивные члены, не содержащие дизъюнкции, т. е. однозначные в синтаксическом отношении. Значение функции истинности для них поэтому сразу приравнивается единице, например $f(V_{sing}) = 1$. (В дальнейшем в таких случаях для упрощения записи будем просто писать $V_{sing} = 1$).

Производимые с полученным выражением действия распадаются на несколько этапов.

Этап 1. Произведение операции переноса.

Имеем $\vec{u}\left(\frac{D}{S_1 + S_2}\right) \times S = S_2^D$.

Второй оператор переноса $(\vec{u})\frac{N+A}{S_2}$ не работает, так как нет S_2, он опускается; то же относится к третьему оператору (к слову *ihr*), причем здесь опускается и связанный с ним конъюнкцией элемент A^S (конъюнкция, один из элементов которой равен нулю, равна нулю). Не работает и первый оператор условного переноса.

Второй оператор условного переноса дает:

$$\left[\vec{A}^S \times b\vec{u}\left(\frac{A}{S_1} + \frac{D}{S_{pl}}\right) + \vec{A}^S\right] \times S_{pl} = A^S \times S_{pl}^D.$$

После работы оператора переноса в S, на которые был произведен перенос, опускаются индексы рода (1, 2, 3), остается только индекс множественности.

Итак, наше выражение приняло вид:

$$\left(S_{pl}^{N+A} + \overrightarrow{S_{sing}^{N+A}}\right) \times V_{sing} \times {}^{(i)}S_{sing}^{N+A} \times (\vec{pr} + \vec{T}^V) \times S_2^D \times \left[\vec{Q}_S\left(\frac{A}{S_2}\right) \times\right.$$

$$\left. \times S_{sing}^{N+A}\right] \times A^V \times V_{sing} \times A^V \times I''(V \,\&\, A^V + A^S) \times (\vec{pr} + \vec{T}^V) \times A^S + S_{pl}^D.$$

Этап 2. Начиная с конца предложения (установление конца производится на основании специальных правил, а мы могли бы лишь проверить, правильно ли он установлен), производим объединения на основании правил (10), и, если это возможно (ср. § 8), применяется оператор *pr*. Начиная с этого этапа, на каждом шаге отмечается истинность соответствующих элементов.

2а. $\vec{A}^S \times S_{pl} = S_{pl}$ и $A^S = 1$ (по (10)).

2б. $(\vec{pr}+\vec{T}^V)\times S_{pl}=(\vec{pr}\times S_{pl})+(\vec{T}^V\times S_{pl})$; конъюнкция во второй части равна нулю (по первому правилу установления истинности элементов), вторая часть отбрасывается, первая часть может быть объединена с элементом под оператором I'' (по определению этого оператора). Имеем

$V\times(\vec{pr}\times S_{pl})=V$ и $pr=1$ и $S_{pl}=1$ (по эквивалентности (11)).

2в. $A^V\times V=V$ и $A^V=1$ (по определению I'' и (10)).

2г. $A^V\times V_{sing}=V_{siug}$ и $A^V=1$ (по (10)).

Мы получили выражение:

$$\left(S_{pl}^{N+A}+S_{sing}^{\overrightarrow{N+A}}\right)\times V_{sing}\times{}^{(i)}S_{sing}^{N+A}\times(\vec{pr}+\vec{T}^V)\times\left[\vec{Q}_S\left(\frac{A}{S_2}\right)\times S_{sing}^{N+A}+S_{sing}^{N+A}\right]\times$$
$$\times V_{sing}\times I''(V\&A^V+A^S)$$

и ряд истинных элементов.

Этап 3. На третьем этапе вступают в работу другие операции Q и I (вернее — дизъюнкции, содержащие эти операторы). В нашем случае на основании правил в § 8 первым должен применяться оператор Q, так как он находится между ${}^{(i)}S$ и I''.

Итак, возьмем часть $[Q_S(\ )\times S_{sing}^{N+A}+S^{N+A}]\times V_{sing}$.

Это равно по (7)

$$[\vec{Q}_S(\ )\times S_{sing}^{N+A}\times V_{sing}]+[S^{N+A}\times V_{sing}].$$

Обе части этой дизъюнкции можно предположить истинными. Действительно, первая дает Q_S Connex $(S,\ V)$, а вторая Connex $(S,\ V)$ (по (12)).

Однако, если мы предположим истинной вторую, то, как легко убедиться простой проверкой, мы после дальнейших преобразований придем к формуле

$$\text{Connex}\ (S,\ V)\times S^A\times\text{Connex}\ (S,\ V),$$

которая не предусмотрена никакими эквивалентностями и поэтому не правильна. Поэтому считаем, что $[S^{N+A}\times V_{sing}]$ содержит один ложный член, а именно S^{N+A} (V_{sing} истинен с самого начала, так как он однозначен). Истинной считаем часть

$$\vec{Q}(\ )\times S_{sing}^{N+A}\times V_{sing}=Q_S\,\text{Connex}\,(S,\ V)\ \text{и}\ S_{sing}^{N}=1\qquad(\text{по (13)}).$$

$S_2\times\vec{Q}_S\left(\frac{A}{S_2}\right)$ преобразуем в S (по определению оператора Q и (10)) и $Q\left(\frac{A}{S_2}\right)=1$.

Выражение теперь имеет вид:

$$(S_{pl}^{N+A}+S_{sing}^{N+A})\times V_{sing}^{(i)}S_{sing}^{N+A}\times(\vec{pr}+T^V)\times S_2^D\times I''(V\&A^V+A^S).$$

Теперь преобразуем $(pr+T^V)\times S_2^D$ так же, как 2б, получаем, беря V из-под знака оператора I'', просто V и $pr=1$, $S_2^0=1$.

После этого, поскольку к I'' ничего в нашем выражении относиться уже больше не может, $I''(V\&A^V+A^S)$ может быть преобразовано в A^V+A^S и затем объединено с ${}^{(i)}S_{sing}^{N+A}$ и $I''(V\&A^S)=1$.

Остается $(S_{pl}^{N+A}+S_{sing}^{N+A})\times V_{sing}\times S_{sing}^{N+A}$.

Это выражение по (13) можно привести к виду $S_{sing}^{N+A}\times V_{sing}\times S_{sing}^{N+A}$ и дальше к $S_{sing}^{N}\times V_{sing}\times V_{sing}^{A}$, что соответствует (12).

Таким образом, $S_{sing}^{N}=1$ и $S_{sing}^{A}=1$.

Оставив теперь в исходном выражении только истинные значения, получаем

Sie	*liebte*	*es,*	*vor*	*einem*	*Bild,*
S^N_{sing}	V_{sing}	$^{(i)}S^A_{sing}$	*pr*	S^D_{sing}	

das	*ihr*	*gefiel,*	*lange*	*stehenzubleiben*	*mit*
$Q^s\left(\frac{A}{S_n}\right)$	A^V	V_{sing}	A^V	$I''(V \& A^S)$	*pr*

halbgeschlossenen	*Augen.*
A^S	S_{pl}

Синтаксический анализ произведен.

Построенная система, конечно, не совершенна, но она, как представляется, обладает достаточной гибкостью, чтобы выразить предложенной символикой или введением некоторых дополнительных символов все основные синтаксические отношения в языке.

A SIMBOLIZATION OF GERMAN SYNTAX AND ITS APPLICATION TO PROBLEMS OF MECHANICAL TRANSLATION

Summary

A slightly modified form of propositional calculus (the truth value of an element depending on the operations applied to it) is used to provide an automatic procedure for recovering grammatical homonymity. A German sentence serves as an example. The procedure can be applied for mechanical translation from German.

П. А. Соболева

ОПЫТ ИСЧИСЛЕНИЯ ТРАНСФОРМАЦИЙ НА ОСНОВЕ ТЕОРИИ С. К. ШАУМЯНА О ПОРОЖДЕНИИ КЛАССОВ СЛОВ В ПРОЦЕССЕ ПОРОЖДЕНИЯ ГРАММАТИКИ

Прежде чем излагать разработанную нами методику для исчисления трансформаций, остановимся вкратце на порождающей модели С. К. Шаумяна[1].

В отличие от порождающей модели Н. Хомского модель С. К. Шаумяна учитывает не только синтагматическую, но и парадигматическую ось языка, так как теория классов слов строится в качестве органического дополнения к теории порождающего процесса.

Согласно данной теории, каждый ядерный класс слов занимает определенную точку ветвления в дереве порождающего процесса. Формальные параметры классов слов возникают в процессе порождения предложений.

Между классами слов допускается только одно отношение — отношение доминации, выражающееся формулой:

$$((D_1 \ll A) \ll N) \ll (V \gg D_2),$$

которую можно прочесть следующим образом. Глагол (V) есть абсолютный доминирующий член предложения; существительное (N) есть относительный доминирующий член первого ранга; прилагательное (A) есть относительный доминирующий член второго ранга; наречие (D) есть абсолютный доминируемый член предложения.

Данные характеристики классов слов (способность доминировать и быть доминируемым) С. К. Шаумян называет ядерными доминативными параметрами и обозначает следующими индексами:

R_1 — доминативный параметр глагола,

R_2 — доминативный параметр существительного,

R_3 — доминативный параметр прилагательного,

R_{4a} — доминативный параметр наречия, доминируемого глаголом,

R_{4b} — доминативный параметр наречия, доминируемого прилагательным.

Деривационные классы слов порождаются посредством операции отображения ядерных классов через отношения. Данная операция определяется С. К. Шаумяном в терминах математической логики. Если рассматривать доминативные параметры слов как некоторые отношения R_1, R_2, R_3 и т. д., то можно получить R-образы лингви-

[1] См. С. К. Шаумян. Трансформационная грамматика и теория классов слов. Тезисы докладов на конференции по структурной лингвистике, посвященной проблеме трансформационного метода. М., 1961.

стических классов слов N, V, A, D_1, D_2, например R''_2V (глагол в параметре существительного) — отглагольное имя, инфинитив, герундий; R''_3V (глагол в параметре прилагательного) — причастие, отглагольное прилагательное; $R''_{4a}V$ (глагол в параметре наречия) — деепричастие; R''_2A (прилагательное в параметре существительного) — субстантивированное прилагательное, отадъективное имя; $R''_{4a}A$ (прилагательное в параметре наречия) — наречие, образованное от прилагательного; R''_1N (существительное в параметре глагола) — отыменный глагол или присвязочное существительное; $R''_{4a}N$ (существительное в параметре наречия) — существительное в косвенном падеже, управляемое глаголом и т. д.

К деривационным классам можно повторно применить операцию отображения и получить новые деривационные классы второй, третьей и т. д. ступеней производности. Например, $R''_1(R''_2V)$ — присвязочное отглагольное имя или глагол, образованный от отглагольного имени (*учить — учитель — учительствовать* или *быть учителем*), $R''_2(R''_3N)$ — существительное, образованное от отыменного прилагательного (*книга — книжный — книжник*) и т. д.

Если условно обозначить деривационные классы, полученные от любых ядерных на любом деривационном шаге, через X, Y, Z, U, W, то общая формула доминации для классов, полученных на следующем деривационном шаге, будет выглядеть как $((R_{4b}X \ll R_3Y) \ll R_2Z) \ll (R_1U \gg R_{4a}W)$ — для упрощения записи мы опускаем штрихи.

Применив определенные правила развертывания, С. К. Шаумян получает восемь ядерных цепочек, каждая из которых содержит только ядерные классы слов. Приводим полностью список ядерных цепочек:

1) $V + D_2$,
2) V,
3) $N + V + D_2$,
4) $N + V$,
5) $D_1 + A + N + V + D_2$,
6) $D_1 + A + N + V$,
7) $A + N + V + D_2$,
8) $A + N + V$.

От этих ядерных цепочек С. К. Шаумян получает новые производные цепочки с помощью операции аппликации, сущность которой заключается в том, что к любому члену ядерной цепочки можно подключить деривационный класс, непосредственно связанный с ним отношением доминации. При этом первый член доминирует над подключаемым. Например, в цепочке $D_1 + A + N + V + D_2$ к члену V можно подключить производный класс $R_{4a}N$ и получить производную цепочку $D_1 + A + N + V + R_{4a}N + D_2$; или к члену N подключить производный класс R_3V и получить цепочку $D_1 + A + R_3V + N + V + D_2$.

Однако построение грамматики языка с помощью универсальной операции аппликации при данном наборе ядерных элементов (т. е. ядерных классов и ядерных предложений) и при данной формуле доминации наталкивается на ряд трудностей, одной из которых, в частности, является деривация предложений связочного типа, например $N + R_iA$ (*Море было спокойным*) или $N + R_1N$ (*Он студент*). Возможный путь преодоления этих трудностей мы видим во введении операции субституции и в связи с этим в некотором изменении трактовки операции аппликации.

Остановимся подробно на нашем понимании операций аппликации и субституции. Кроме того, введем еще две операции, с помощью которых в рамках данной модели мы считаем возможным получать деривационные предложения от ядерных. Это трансформация обращения и трансформация цепного типа.

Рассмотрим все четыре операции с точки зрения а) состава символов в деривационной цепочке; б) количества символов в деривацион-

ной цепочке; в) отношений доминации в деривационной цепочке; г) обусловленности операций, производимых над классами, подключаемыми к цепочке.

При построении деривационных цепочек будем иметь в виду следующее правило, которое назовем правилом непосредственного следования классов слов в деривационной цепочке. В деривационной цепочке непосредственно следовать друг за другом могут: а) слова одного и того же доминативного параметра, например $\ldots A + R_3N \ldots$, $\ldots R_{4a}N + R_{4a}V \ldots$ и б) слова, непосредственно связанные отношениями доминации, например $\ldots R_3D_2 + N \ldots$, $\ldots R_2A + R_1N \ldots$ (с этой точки зрения цепочка, скажем, типа $\ldots R_2V + D_2 \ldots$ или $R_1(R_2V) + R_3D_2$ была бы невозможна).

ОПЕРАЦИЯ АППЛИКАЦИИ

С помощью операции аппликации будем получать деривационные цепочки символов, подключая к любому классу ядерной цепочки новый класс или новые классы слов того же доминативного параметра.

Возьмем, например, ядерную цепочку типа $A + N + V$ (*Старый сад зацвел*)[2] и присоединим к ядерному классу A деривационный класс R_3N. Тогда получим цепочку типа $[A + R_3N] + N + V$ (*Старый вишневый сад зацвел*) со сложным членом[3] $[A + R_3N]$. Если же к ядерному классу A подключить не один, а сразу два деривационных класса, например R_3N и R_3V, то получится деривационная цепочка типа $[A + R_3V + R_3N] + N + V$ (*Старый запущенный вишневый сад зацвел*). К цепочке $N + V + D_2$ (*Мальчик читает здесь*) можно подключить класс $R_{4a}N$ и получить $N + V + [D_2 + R_{4a}N]$ (*Здесь мальчик читает книгу*). Можно подключить к цепочке и непроизводные классы, например: $A + [N + N] + V + D_2$ (*Там растут красивые деревья и цветы*) и т. д. К деривационной цепочке, получившейся в результате действия операции аппликации, можно повторно применить эту же операцию и получить новую деривационную цепочку, так как деривация цепочек с помощью этой операции теоретически ничем не ограничена.

Получение производных цепочек путем операции аппликации можно суммировать в виде следующей формулы — для полной ядерной цепочки: $[D_1 + R_{4b}X + R_{4b}Y + R_{4b}\ldots] + [A + R_3X + R_3Y + R_3\ldots] + [N + R_2X + R_2Y + R_2\ldots] + [V + R_1X + R_1Y + R_1\ldots] + [D_2 + R_{4a}X + R_{4a}Y + R_{4a}\ldots]$, где X, Y и т. д. — любая ступень производности от любого ядерного класса, включая нулевую. Если X, Y и т. д. — нулевая ступень производности от первого члена соответствующей скобки, то подключаемый класс — непроизводный, так как класс в своем параметре равен самому себе: $R_{4b}D_1 = D_1$; $R_3A = A$; $R_2N = N$; $R_1V = V$; $R_{4a}D_2 = D_2$.

Аналогичные формулы можно было бы написать и для остальных ядерных цепочек, например для двучленной:

$$[V + R_1X + R_1Y + R_1\ldots] + [D_2 + R_{4a}X + R_{4a}Y + R_{4a}\ldots]$$

и т. д.

Из данных формул следует, что если операция аппликации производится над любой ядерной цепочкой, кроме полной, состав символов,

---

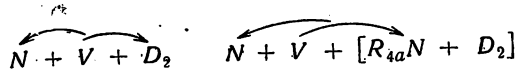
[2] Лингвистическая интерпретация цепочек классов слов должна пониматься как одна из возможных, в частности для русского языка. При этом не имеется в виду прямое соответствие между порядком символов в цепочке и порядком слов в иллюстративном предложении. Так, например, данную цепочку можно интепретировать и как *Зацвел старый сад*.

[3] Подробнее о понятии сложного члена см. стр. 241.

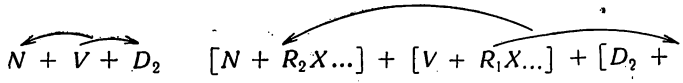
т. е. набор исходных ядерных классов, в ней может меняться, например: $N + V + D_2 \rightarrow [N + R_2A] + V + D_2$ или $N + V \rightarrow [N + R_2D_2] + V$.

Из формул также следует, что если к цепочке подключается более чем один член, то количество и характер шагов по его деривации не обусловлены количеством и характером шагов по деривации предыдущего подключенного члена. Так, например, в цепочке $A + N + V + D_2$ мы можем добавить к члену A класс первой ступени производности типа R_3N, а к члену D_2 добавить класс третьей ступени производности типа $R_{4a}(R_2(R_3V))$ (*Новый учитель работает здесь* → *Новый школьный учитель работает с отстающими здесь*). Добавление одного члена никак не обусловливает добавления другого, так же как операции по получению одного никак не влияют на операции по получению другого.

Отношения доминации в деривационной цепочке, полученной способом аппликации, изоморфны отношениям доминации в исходной ядерной. Из-за того, что каждый подключаемый член имеет тот же доминативный параметр, что и соответствующий ядерный, в деривационной цепочке можно расставить те же «стрелки доминации», что и в ядерной. Например:



или в общем виде:



Обратим внимание на тот факт, что как в ядерной, так и в деривационной цепочке аналогичные стрелки связывают члены, в состав которых входят одни и те же классы (т. е. в исходной цепочке доминирующим членом был V, а доминируемыми — N и D_2, и в производной цепочке в состав доминирующего члена также входит класс V, а в состав доминируемых — классы N и D_2). Последнее мы называем сохранением исходных отношений доминации.

Итак, деривационная цепочка, полученная с помощью операции аппликации, характеризуется тем, что в ней:

а) сохраняются исходные отношения доминации;

б) не обязательно сохраняется исходный состав ядерных классов;

в) не сохраняется исходное количество ядерных классов;

г) операции по деривации подключаемых членов не взаимообусловлены.

ОПЕРАЦИЯ СУБСТИТУЦИИ

С помощью операции субституции можно получать деривационные цепочки классов, заменяя любой класс ядерной цепочки классом того же доминативного параметра. Так, например, если в ядерной цепочке $N + V + D_2$ каждый класс последовательно заменять классами того же доминативного параметра, то получится ряд деривационных цепочек типа:

$R_2V + V + D_2$ *(Там происходит собрание)*,
$R_2A + V + D_2$ *(Иногда случается необычное)*,
$N + R_1N + D_2$ *(Он здесь директором)*,
$N + R_1A + D_2$ *(Она всегда весела)*,
$N + V + R_{4a}N$ *(Охотник убил медведя)*,

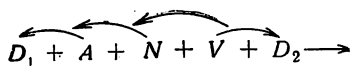
$N + V + R_{4a}A$ *(Птица летит высоко)*,
$N + V + R_{4a}V$ *(Он говорит, захлебываясь)* и т. д.

Таким образом можно получить бесконечное число деривационных цепочек, заменяя ядерные классы слов словами второй, третьей и т. д. ступеней производности, заменяя сразу два или более класса.

Для каждой ядерной цепочки можно было бы написать общую формулу деривации путем субституции, например для полной: $R_{4b}X + R_3Y + R_2Z + R_1U + R_{4a}W$, где $X \ldots W$ — любые ядерные классы любой ступени производности. Или для неполной ядерной цепочки, например: $R_2X + R_1Y$ и т. д.

Из формулы деривации путем субституции следует, что для любой цепочки, кроме полной, сохранение состава ядерных классов — условие необязательное, в то время как сохранение количества ядерных классов — условие обязательное.

О сохранении исходных отношений доминации можно говорить лишь в частном случае цепочек типа





где аналогичные стрелки связывают члены, содержащие одни и те же ядерные классы. В остальных же случаях исходные отношения доминации не сохраняются, так как характерным для операции субституции как раз является замена классов.

Из формулы субституции также следует, что количество и характер операций над деривационными классами, подключаемыми к цепочке, вполне произволен, так же как и выбор самого деривационного класса. Обязателен лишь выбор последнего параметра деривационного класса.

Итак, деривационная цепочка, полученная с помощью операции субституции, характеризуется тем, что в ней:

а) сохраняется исходное количество ядерных классов, но не обязательно сохраняется их состав;

б) не обязательно сохраняются исходные отношения доминации:

в) операции по получению подключаемых членов не взаимообусловлены.

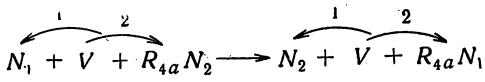
ТРАНСФОРМАЦИЯ ОБРАЩЕНИЯ

Трансформация обращения предусмотрена прежде всего для деривации новых цепочек от неядерных. В этом случае она применима к цепочкам, в составе которых имеется минимум два члена, содержащих один и тот же символ, например к цепочкам типа $N_1 + V + R_{4a}N_2$ *(Рабочие строят дом)* или $R_3V_1 + N + V_2$ *(Играющий мальчик смеется)* или $N + V_1 + R_{4a}V_2$ *(Он учится, работая)* и т. д. Сущность операции обращения заключается во взаимной перестановке (взаимном изменении параметров) одноименных символов: $N_1 + V + R_{4a}N_2 \rightarrow N_2 + V + R_{4a}N_1$ *(Рабочие строят дом → Дом строится рабочими* или *Книга имеет недостатки → В книге имеются недостатки)*; $R_3V_1 + N + V_2 \rightarrow R_3V_2 + N + V_1$ *(Играющий мальчик смеется → Смеющийся мальчик играет)*; $N + V_1 + R_{4a}V_2 \rightarrow N + V_2 + R_{4a}V_1$ *(Он учится, работая → Он работает, учась)*.

Объяснительная сила этой операции может возрасти, если допустить взаимную перестановку положительных и нулевых символов. В таком случае она будет применима и к ядерным цепочкам. Обозначим нулем отсутствие символа в каком-либо параметре, например в параметре прилагательного: $0 + N + V$. Обменяем параметрами символы V и 0. Перевод V на место нуля дает член R_3V, интерпретируемый как причастие. Перевод нуля в параметр глагола дает R_10 или $0:0 + N + V \rightarrow R_3V + N + 0$ (*Человек скучает* → *Скучающий человек*).

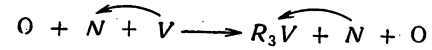
Данная операция, заключающаяся во взаимном изменении доминативных параметров классов слов (взаимной перестановке символов), влечет за собой изменение исходных отношений доминации.

Для того, чтобы разобраться, в чем заключается изменение исходных отношений доминации при деривации цепочек путем обращения, обозначим в первом примере доминацию группы глагола над N_1 доминативной связью 1, а доминацию глагола над словом, находящимся в параметре наречия, доминативной связью 2. Тогда, пронумеровав стрелки доминации в исходной и производной цепочках —



увидим, что «направление» этих связей изменилось на обратное. Стрелка 1, связывавшая в исходной цепочке группу глагола и N_1, в производной связывает группу глагола и N_2, а стрелка 2 в исходной цепочке, связывавшая классы V и N_2, в производной связывает V и N_1.

Еще нагляднее изменение отношений доминации видно на примерах типа



(*Человек скучает* → *Скучающий человек*). В исходной цепочке V доминирует над N, а в производной N доминирует над членом, содержащим символ V.

При трансформации обращения заслуживает внимания тот факт, что деривационные шаги, необходимые для получения подключаемых классов, взаимообусловлены. Так, например, мы переводим класс N_1 в параметр R_{4a} потому и только потому, что класс N_2 переводится в параметр R_2 и т. д.

Итак, деривационная цепочка, полученная с помощью трансформации обращения, характеризуется тем, что в ней:

а) обязательно сохраняется исходный состав и количество ядерных классов;

б) обязательно меняются исходные отношения доминации;

в) операции над подключаемыми классами взаимообусловлены.

(Переход от модели к реальному языку наложит ряд ограничений на деривацию цепочек с помощью трансформации обращения, в частности для того, чтобы исключить соотношения типа *Жена брата приехала* → *Брат жены приехал* или *Охотник убил медведя* → *Медведь убил охотника*. В ряде случаев можно будет воспользоваться критерием нетождественности конструкций, предложенным Ю. Д. Апресяном[4]).

Операции аппликации, субституции и трансформация обращения обладают определенной объяснительной силой для языка, однако с помощью правил, сводящихся к тому, чтобы добавлять или заме-

[4] Ю. Д. Апресян. Метод непосредственно составляющих и трансформационный метод в структурной лингвистике. «Русский язык в национальной школе», 1962, № 4.

щать некоторые классы классами того же доминативного параметра, невозможно объяснить целый ряд предложений, носящих, интуитивно, явно деривационный характер, например таких, как... *его здешняя работа*... или *Он здешний рабочий*, понимаемых как производные от цепочки: *Он работает здесь*.

В самом деле, если переписать данные три предложения в терминах, применявшихся выше, то получим следующее:

1) $N + V + D_2$ (*Он работает здесь*),
2) $[R_3N + R_3D_2] + R_2V$ (*Его здешняя работа*),
3) $N + R_1(R_2V + R_3D_2)$[5] (*Он здешний рабочий*).

Сравнение цепочек показывает, что если в первой отсутствует класс параметра R_3, то во второй и третьей содержатся таковые. Во второй цепочке нет класса параметра R_1 (т. е. она не является предложением), а во второй и третьей нет класса параметра R_{4a}, хотя в исходной цепочке имеется как класс параметра R_1, так и класс параметра R_{4a}.

Очевидно, порождение второй и третьей цепочки не может быть объяснено правилами аппликации, субституции и обращения в том виде, в котором они сформулированы выше. Тем не менее легко заметить, что деривационные цепочки данного типа характеризуются тем же составом и количеством ядерных классов, что и соответствующая исходная, и аналогичными отношениями доминации. Так, в обеих цепочках доминирует член, содержащий класс V, а доминируемыми являются члены, содержащие N и D_2. Кроме того, если в данных предложениях расставить доминативные связи, то мы увидим, что их направления не изменились — они направлены на те же классы:

$$N + V + D;$$

$$[R_3N + R_3D_2] + R_2V$$

или, если раскрыть скобки:

$$R_3N + R_3D_2 + R_2V$$

$$N + R_1(R_2V + R_3D_2).$$

Необходимое правило, которое нужно ввести для деривации цепочек такого типа, заключается в обусловленности операций, проводимых над доминируемыми членами, операциями над доминирующим членом.

Условимся называть деривацию, при которой у производной цепочки а) сохраняется состав ядерных классов исходной; б) сохраняется количество ядерных классов исходной; в) сохраняются исходные отношения доминации; г) операции над доминируемыми членами обусловлены операциями над доминирующим, трансформацией цепного типа[6].

---

[5] Порядок заключения в скобки, вытекающий из правила непосредственного следования слов в деривационной цепочке, объясняется ниже. См. стр. 241.

[6] Ср. более раннее определение трансформации, приведенное в докладе «О трансформационном анализе словообразовательных отношений», прочитанном на конференции по структурной лингвистике, посвященной проблемам трансформационного

Название отражает тот факт, что преобразование одного члена цепочки автоматически влечет за собой преобразование других ее членов, что в свою очередь может привести к преобразованию цепочки в целом.

ТРАНСФОРМАЦИЯ ЦЕПНОГО ТИПА

Условие сохранения прежнего состава ядерных классов, исходных отношений доминации и обусловленности операций над доминируемыми членами диктует следующую процедуру получения трансформов данного типа:

1. Над доминирующим членом ядерной цепочки производится операция или ряд операций отображения, в результате чего от него образуется некоторый деривационный класс.

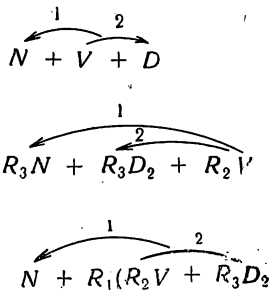
2. Полученный деривационный класс подключается к новой цепочке в соответствии с его доминативным параметром.

3. Над ядерным классом, непосредственно связанным с первым отношениями доминации, производится операция отображения (положительная или нулевая), обусловленная новым доминативным параметром предыдущего класса, т. е., иными словами, от другого ядерного класса получается деривационный класс, доминативный параметр которого обусловлен доминативным параметром предыдущего, и т. д. до тех пор, пока все символы из ядерной цепочки ни будут перенесены в деривационную.

Таким образом, сущность цепной трансформации состоит в составлении новых цепочек из членов старой, над которыми произведены некоторые обусловленные операции.

Доминативные свойства глагола. Выше говорилось о том, что доминативный параметр каждого последующего члена в трансформе обусловлен доминативным параметром предыдущего и что строится такой трансформ, начиная с доминирующего члена. А так как абсолютным доминирующим членом любого ядерного предложения является глагол, остановимся подробно на его доминативных свойствах и на доминативных свойствах его дериватов.

На стр. 238 мы обозначили свойство глагола доминировать над словом в параметре наречия доминативной связью 2, а свойство группы глагола доминировать над словом в параметре существительного — доминативной связью 1. Аналогичным образом мы назвали и связи между соответствующими производными от этих классов, например:



Введем правило реализации доминативных связей отглагольного деривата в трансформе.

метода (декабрь 1961 г.). Трансформация определялась нами как такое преобразование синтагмы, которое является следствием операции, произведенной над доминирующим членом, и при котором сохраняются исходные отношения доминации. См. «Известия ОЛЯ АН СССР», т. XXI, вып. 2, 1962.

1. Доминация над классом, связанным с отглагольным дериватом первой доминативной связью, осуществляется по последнему параметру (последнему деривационному шагу) этого деривата.

2. Доминация над классом, связанным с отглагольным дериватом второй доминативной связью, осуществляется:

а) по последнему параметру (последнему деривационному шагу) этого деривата;

б) по предпоследнему параметру (предпоследнему деривационному шагу) этого деривата.

Во втором случае предполагается объединение доминирующего и доминируемого члена в один сложный член, общий доминативный параметр которого определяется последним деривационным шагом отглагольного деривата. При записи цепочки, получившейся в результате применения правила 2б, доминируемый член подключается к доминирующему скобочно. В противном случае получались бы невозможные цепочки, противоречащие правилу «непосредственного следования классов слов» (см. стр. 235). Иными словами, правило 2б предполагает возможность переключения цепочек на уровень классов. (Лингвистический смысл правила 2б заключается в том, что «производящая основа» может сохранять свои фразовые связи, свое «управление». Например: *руководить работой → руководство работой*, *стремиться к счастью → стремление к счастью*).

Поясним оба правила. Если, например, доминирующему члену из цепочки $N+V+D_2$ задать значение R_2V, т. е. перевести его в параметр существительного, то, по условиям 1 и 2а, члены N и D_2 примут строго определенные значения, обусловленные параметром R_2, точнее — его свойством доминировать только над словами параметра R_3. Именно поэтому производные от N и D_2 примут вид R_3N и R_3D_2, а вся цепочка — $[R_3N+R_3D_2]+R_2V$ (*он работает здесь → . . .его здешняя работа . . .*). В данном случае перевод V в параметр R_2 вызвал перевод N и D_2 в параметр R_3.

Однако $[R_3N+R_3D_2]+R_2V$ — не единственный трансформ цепочки $N+V+D_2$ при заданном превращении V в R_2V. По правилу 1 и 2б она может принять вид: $R_3N+R_2(V+D_2)$, так как предпоследним доминативным параметром члена R_2V является параметр V, который предполагает возможность непосредственной доминации над наречием (*Он работает здесь → Его работа здесь*).

Из данного примера также видно, что частным случаем зависимости операций, проводимых над доминируемым членом, от операций над доминирующим является нулевая операция над первым.

Разберем еще один пример. Предположим, что члену V из ядерной цепочки $N+V+D_2$ задается значение $R_1(R_2V)$. Тогда по правилу 1 и 2а ее трансформ примет вид $N+R_1(R_2V)+D_2$, так как параметр R_1 предполагает непосредственную доминацию над N и D_2. В данном случае результат трансформации равносилен результату субституции класса V классом $R_1(R_2V)$ (*Он преподавал тогда → Он был преподавателем тогда*)[7]. В то же время по правилам 1 и 2б трансформ имеет вид $N+R_1(R_2V+R_3D_2)$, так как предпоследним параметром доминирующего члена является R_2, а он предполагает непосредственную доминацию только над классом параметра R_3 (*Он преподает здесь → Он здешний преподаватель*).

---

[7] Такое перекрещивание результатов различных операций в принципе является нежелательным. Однако в тех случаях, где это неизбежно, важно строго определить условия такого перекрещивания. Так, например, в нашей системе результат трансформации равносилен результату субституции только в тех случаях, когда последний деривационный шаг доминирующего члена переводит его в тот же параметр, в котором он находился в исходной цепочке. Это создает условия для перевода доминируемых членов в новую цепочку без всякого изменения.

В связи с возможностью скобочного подключения доминируемых классов по правилу 2б уместно уточнить понятие члена цепочки.

Если ядерный или деривационный класс, отделенный от другого хотя бы одним знаком +, назвать простым членом цепочки (например $\ldots+N+\ldots$ или $\ldots+R_3D_2$ или $\ldots+R_1(R_2V+\ldots)$, то группу простых классов, входящих в один и тот же доминативный параметр, можно назвать сложным членом цепочки. Сложными членами тогда будут такие, как $\ldots+R_1(V+D_2)$ или $\ldots+R_1(R_2V+R_3D_2)$, а также $\ldots+[A+R_3N]+\ldots$ или $\ldots+[V+V]+\ldots$, полученные в результате операции аппликации.

Подведем итог рассмотренным выше типам преобразований. Все четыре типа преобразований — аппликация, субституция, трансформация обращения и цепная трансформация — рассматривались нами с точки зрения: а) состава ядерных классов, б) количества ядерных классов, в) отношений доминации, г) обусловленности операций над членами деривационной цепочки.

Если все эти признаки рассматривать как дифференциальные признаки деривации, то каждый тип деривации можно будет определить через совокупность тех или иных признаков (см. табл. 1).

Таблица 1

Дифференциальный признак \ Тип деривации	Аппликация	Субституция	Трансформация обращения	Трансформация цепная
Сохранение состава ядерных классов	0	0	+	+
Сохранение количества ядерных классов	—	+	+	+
Сохранение исходных отношений доминации	+	0	—	+
Обусловленность операций над доминируемыми членами . . .	—	—	+	+

Знак + обозначает «да», знак — обозначает «нет», а 0 обозначает «возможно, но не обязательно».

ИСЧИСЛЕНИЕ ТРАНСФОРМАЦИЙ ЦЕПНОГО ТИПА

Зная правила деривации трансформов цепного типа, поставим перед собой задачу получить полный список таких трансформов от ряда произвольно выбранных цепочек.

В качестве исходных цепочек возьмем ядерную цепочку $N+V+D_2$, три возможные цепочки, полученные от первой путем субституции последнего члена классами первой ступени производности от всех остальных ядерных классов (кроме D_1, так как класс $R_{4a}D_1$ имеет тот же лингвистический смысл, что и $R_{4a}D_2$, т. е. D_2): $N+V+R_{4a}A$; $N+V_1+R_{4a}V_2$; $N_1+V+R_{4a}N_2$, и одну цепочку, полученную от последней путем трансформации обращения: $N_2+V+R_{4a}N_1$.

Ограничим свою задачу получением трансформов, где доминирующему члену последовательно задаются значения R_2V и $R_1(R_2V)$, т. е. где глагол обращается в отглагольное имя и присвязочное отглагольное имя.

Лингвистический смысл данной работы заключается в том, чтобы получить максимально полный список трансформаций, связывающих синтагмы, где доминирует глагол, с синтагмами, где в качестве домини-

рующего члена выступает отглагольное имя, или, иными словами, получить трансформации, связывающие глагол и отглагольное имя на синтаксическом уровне.

Выбор в качестве исходных именно данных цепочек символов обусловлен тем, что они с достаточной полнотой отражают свойства глагола сочетаться с наречием, существительным в косвенном падеже (с предлогом и без предлога), с инфинитивом и деепричастием, а также с существительным в именительном падеже, выполняющим функцию подлежащего при этом глаголе.

Полный список трансформаций для такого рода синтагм покажет, каким возможным преобразованием подвергаются ее члены при заданном обращении глагола в отглагольное имя.

1. Трансформы цепочки $N+V+D_2$. Выше, при рассмотрении правил цепных трансформаций (1, 2а, 2б) мы уже вывели четыре возможных трансформа $N+V+D_2$:

а) $[R_3N+R_3D_2]+R_2V$;

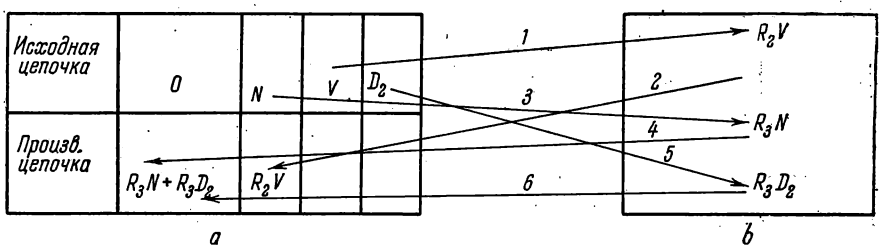
б) $R_3N+R_2(V+D_2)$;

в) $N+R_1(R_2V)+D_2$;

г) $N+R_1(R_2V+R_3D_2)$.

Для того, чтобы более наглядно представить себе все шаги, необходимые для деривации этих цепочек, воспользуемся моделью порождающего устройства со спаренными генераторами, предложенной С. К. Шаумяном. Модель состоит из трех основных частей: 1) источника начальной информации; 2) генератора цепочек символов; 3) генератора классов символов. Между генераторами существует обратная связь. Работа генераторов моделирует взаимодействие синтагматической и парадигматической осей языка [8].

Схема работы генераторов при получении цепочки «а»



a. Генератор цепочек символов *b*. Генератор классов символов

Такт 1. Перевод V в генератор классов для получения деривационного класса R_2V.

Такт 2. Помещение его в производную цепочку на место, обусловленное параметром R_2.

Такт 3. Перевод N в генератор классов и получение R_3N по правилу 1.

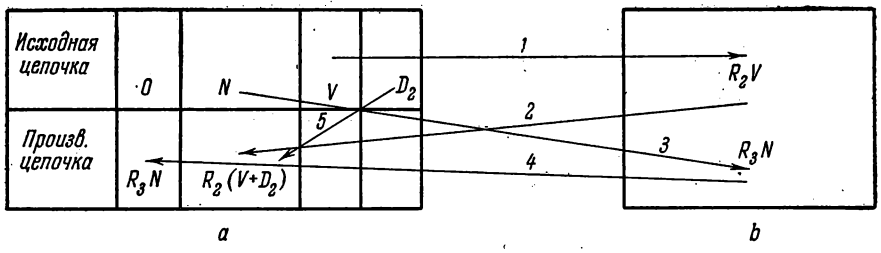
Такт 4. Помещение R_3N в производную цепочку на место, обусловленное его доминативным параметром.

Такт 5. Перевод D_2 в генератор классов и получение деривационного класса R_3D_2 по правилу 2а.

Такт 6. Помещение члена R_3D_2 в производную цепочку на место, обусловленное его параметром. Поскольку в параметре R_3 уже находится некоторый класс, то R_3D_2 подключается к нему с помощью операции аппликации.

[8] Подробнее см.: С. К. Шаумян. Насущные задачи структурной лингвистики. «Известия ОЛЯ АН СССР», т. XXI, вып. 2, 1962.

Схема работы генераторов при получении цепочки «б»



a. Генератор цепочек символов *b*. Генератор классов символов

Такт 1. Перевод V в генератор классов для получения производного класса R_2V.

Такт 2. Помещение его в производную цепочку на место, обусловленное параметром R_2.

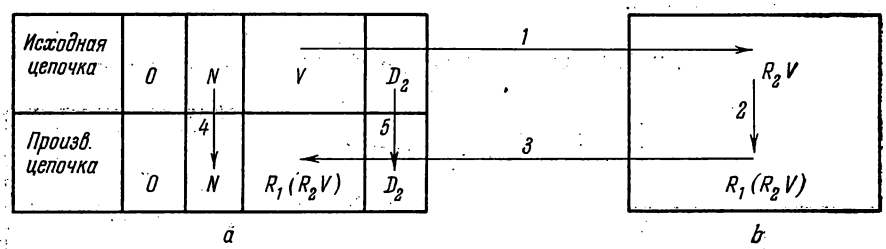
Такт 3. Перевод N в генератор классов и получение R_3N по правилу 1.

Такт 4. Помещение класса R_3N в деривационную цепочку на место, обусловленное его доминативным параметром.

Такт 5. Перевод D_2 непосредственно из ядерной цепочки в деривационную и подключение его к V по правилу 2б.

Как следует из схемы, различие в работе генераторов при получении цепочек «а» и «б» проявляется только начиная с такта 5. Первые четыре такта в обоих случаях совпадают.

Схема работы генераторов при получении цепочки «в»



a. Генератор цепочек символов *b*. Генератор классов символов

Такт 1. Перевод V в генератор классов для получения R_2V.

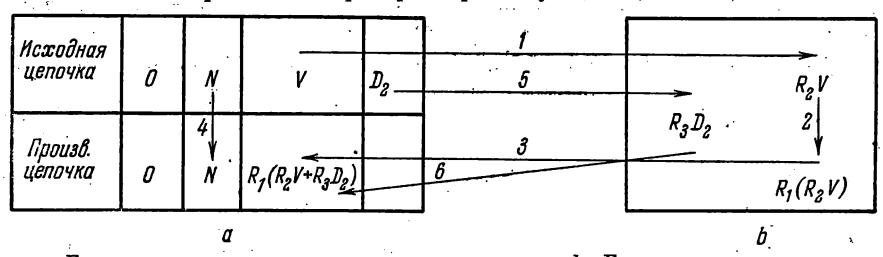
Такт 2. Получение от R_2V производного класса второй ступени $R_1(R_2V)$ путем отображения его через параметр глагола.

Такт 3. Помещение $R_1(R_2V)$ в производную цепочку на место, обусловленное параметром R_1.

Такт 4. Перевод N непосредственно в производную цепочку по правилу 1.

Такт 5. Перевод D_2 непосредственно в производную цепочку по правилу 2а.

Схема работы генераторов при получении цепочки «г»



a. Генератор цепочек символов *b*. Генератор классов символов

Такт 1. Перевод V в генератор классов для получения R_2V.

Такт 2. Получение от R_2V производного класса второй ступени $R_1(R_2V)$.

Такт 3. Помещение $R_1(R_2V)$ в производную цепочку на место, обусловленное параметром R_1.

Такт 4. Перевод N непосредственно в производную цепочку по правилу 1.

Такт 5. Перевод D_2 в генератор классов и получение производного класса R_3D_2, обусловленного предпоследним параметром R_2 класса $R_1(R_2V)$ по правилу 2б.

Такт 6. Подключение класса R_3D_2 непосредственно к R_2V в производной цепочке.

Различие в работе генераторов при получении цепочек «в» и «г» также начинается только с такта 5.

Если теперь сравнить работу генераторов во всех четырех случаях, то мы увидим, что третий и четвертый отличаются от первого и второго начиная с такта 2 (для получения цепочек «в» и «г» над доминирующим членом проводится на одну операцию больше, чем при получении цепочек «а» и «б»).

2. Трансформы цепочки $N+V+R_4A$ [9] (лингвистическая интерпретация такой цепочки — предложения типа *Мальчик бежит быстро*, *Ярко светит солнце*).

Если перевести V в параметр R_2, то, по правилам 1 и 2а, N и R_4A автоматически переводятся в параметр R_3, образуя деривационные классы R_3N и $R_3(R_4A)$. Производная цепочка будет иметь вид:

д) $[R_3N+R_3(R_4A)]+R_2V$ (*Солнце светит ярко* → ... *яркий солнечный свет*...)[10].

Цепочка «д» — не единственный трансформ исходной $N+V+R_4A$ при заданном обращении V в R_2V. Если учесть правило 2б, то трансформ будет иметь вид:

е) $R_3N+R_2(V+R_4A)$ (*Они говорят по-французски* → ... *их разговор по-французски*...).

Если же доминирующему члену исходной цепочки задать значение $R_1(R_2V)$, то по правилу 1 и 2а получим трансформ:

ж) $N+R_1(R_2V)+R_4A$ (*Книга давно печатается* → *Книга давно в печати*).

По правилам 1 и 2б получаем трансформ:

з) $N+R_1(R_2V+R_3(R_4A))$ (*Он метко стреляет* → *Он меткий стрелок*).

Цепочки «д», «е», «ж», «з» являются трансформами исходной цепочки $N+V+R_4A$ при заданных обращениях V в R_2V и в $R_1(R_2V)$. Схемы получения цепочек «д», «е», «ж», «з» аналогичны схемам получения цепочек «а», «б», «в», «г».

3. Трансформы цепочки $N+V_1+R_4V_2$. Не повторяя рассуждений, аналогичных приведенным выше, дадим полный список трансформов цепочки $N+V_1+R_4V_2$ при заданном обращении V в R_2V и в $R_1(R_2V)$.

В первом случае имеем (правило 1 и 2а):

и) $[R_3N+R_3(R_4V_2)]+R_2V_1$ (*Он смеется, захлебываясь* → ... *его захлебывающийся смех*...)

а также (правило 1 и 2б):

---

[9] Здесь и в дальнейшем будем для краткости параметр наречия, доминируемого глаголом, обозначать R_4 вместо R_4a.

[10] Мы не даем всех возможных интерпретаций полученных цепочек. Более полная интерпретация будет дана ниже после установления специальных правил, см. стр. 248—249.

к) $R_3N + R_2(V_1 + R_4V_2)$ (*Он желает учиться → ... его желание учиться...*).

Во втором случае, т. е. при обращении V в $R_1(R_2V)$, имеем (по правилу 1 и 2а):

л) $N + R_1(R_2V_1) + R_4V_2$ (*Он любит пошутить → Он любитель пошутить*),

а также (правило 1 и 2б):

м) $N + R_1(R_2V_1 + R_3(R_4V_2))$ (*Она учится, работая → Она работающая ученица*).

4. Трансформы цепочки $N_1 + V + R_4N_2$. Аналогичным образом получаем трансформы для цепочки $N_1 + V + R_4N_2$. При V, обращающемся в R_2V, имеем (1 и 2а):

н) $[R_3N_1 + R_3(R_4N_2)] + R_2V$ (*Они строят станки → ... их станкостроение...* или *Они разбирают дело → ... разбор дела ими...*).

При том же значении доминирующего члена имеем (1 и 2б):

о) $R_3N_1 + R_2(V + R_4N_2)$ (*Он руководит предприятием → ... его руководство предприятием...*).

Если доминирующий член принимает значение $R_1(R_2V)$, то по тем же правилам получаем два следующих трансформа:

п) $N_1 + R_1(R_2V) + R_4N_2$ (*Он борется за мир → Он борец за мир*);

р) $N_1 + R_1(R_2V + R_3(R_4N_2))$ (*Он водит автобус → Он водитель автобуса* или *Он любит правду → Он правдолюбец*).

5. Трансформ цепочки $N_2 + V + R_4N_1$. Данная цепочка получена в результате трансформации обращения, которой было подвергнуто предложение $N_1 + V + R_4N_2$. При V, переходящем в R_2V, имеем:

с) $[R_3N_2 + R_3(R_4N_1)] + R_2V$ (*Ими приносится жертва → ... их жертвоприношение...*),

а также

т) $R_3N_2 + R_2(V + R_4N_1)$ (*Лекция читается профессором → ... чтение лекции профессором...* или *В совхозе заготавливаются овощи → ... заготовка овощей в совхозе...*).

Если V_1 переходит в $R_1(R_2V)$, то получаем следующие два трансформа:

у) $N_2 + R_1(R_2V) + R_4N_1$ (*Дело рассматривалось в суде → Дело было на рассмотрении в суде*);

ф) $N_2 + R_1(R_2V + R_3(R_4N_1))$ (*Радио изобретено Поповым → Радио — изобретение Попова*).

Перейдем теперь к полной интерпретации полученных нами деривационных цепочек «а»...«ф» на материале русского языка. Для удобства дальнейшего рассмотрения выпишем исходные цепочки слева, а их трансформы справа:

1) $N + V + D_2$:
- а) $[R_3N + R_3D_2] + R_2V$,
- б) $R_3N + R_2(V + D_2)$,
- в) $N + R_1(R_2V) + D_2$,
- г) $N + R_1(R_2V + R_3D_2)$;

2) $N + V + R_4A$:
- д) $[R_3N + R_3(R_4A)] + R_2V$,
- е) $R_3N + R_2(V + R_4A)$,
- ж) $N + R_1(R_2V) + R_4A$,
- з) $N + R_1(R_2V + R_3(R_4A))$;

3) $N + V_1 + R_4V_2$: и) $[R_3N + R_3(R_4V_2)] + R_2V_1$,
к) $R_3N + R_2(V_1 + R_4V_2)$,
л) $N + R_1(R_2V_1) + R_4V_2$,
м) $N + R_1(R_2V_1 + R_3(R_4V_2))$;

4) $N_1 + V + R_4N_2$: н) $[R_3N_1 + R_3(R_4N_2)] + R_2V$,
о) $R_3N_1 + R_2(V + R_4N_2)$,
п) $N_1 + R_1(R_2V) + R_4N_2$,
р) $N_1 + R_1(R_2V + R_3(R_4N_2))$;

5) $N_2 + V + R_4N_1$: с) $[R_3N_2 + R_3(R_4N_1)] + R_2V$,
т) $R_3N_2 + R_2(V + R_4N_1)$,
у) $N_2 + R_1(R_2V) + R_4N_1$,
ф) $N_2 + R_1(R_2V + R_3(R_4N_1))$.

Так как в основе получения деривационных цепочек «а». . .«г»; «д». . . «з»; «и». . .«м»; «н». . .«р» и «с». . .«ф» лежат одни и те же правила, а при записи первого члена как $N_{(1,2)}$ соответствующие им исходные цепочки отличаются друг от друга только значением символа, находящегося в параметре R_4, то полученный список формул можно было бы выразить более компактно, обозначив ядерный класс в параметре R_4 через переменную X, которая может принимать значения D_2, A, V_2, N_1, N_2. Тогда пять исходных цепочек можно записать в виде одной формулы (1—5): $N_{(1,2)}+V+R_4X$, а двадцать производных — в виде следующих четырех формул:

(«а», «д», «и», «н», «с»)	$[R_3N_{(1,2)}+R_3(R_4X)]+R_2V$;
(«б», «е», «к», «о», «т»)	$R_3N_{(1,2)}+R_2(V+R_4X)$;
(«в», «ж», «л», «п», «у»)	$N_{(1,2)}+R_1(R_2V)+R_4X$;
(«г», «з», «м», «р», «ф»)	$N_{(1,2)}+R_1(R_2V+R_3(R_4X))$.

Однако при переписке полученных формул в термины, устанавливаемые правилами интерпретации, удобнее пользоваться первым, т. е. развернутым списком.

Правила интерпретации. Для того, чтобы от абстрактной модели перейти к реальному языку, следует установить правила интерпретации этой абстрактной модели. Правила интерпретации обусловлены спецификой языка или языков, для которых они составляются, а также тем, составляются ли они для грамматики языка в целом или для некоторого фрагмента этой грамматики.

Поясним данные положения. Правила интерпретации модели различны для разных языков потому, что грамматические значения получают в них различное формальное выражение. Так, например, если бы мы составляли правила интерпретации для описания грамматики английского языка, то символ R_2V следовало бы интерпретировать как отглагольное имя, инфинитив и герундий. Для русского же языка данный символ получает только две первые интерпретации. Или символ R_3N в русском языке интерпретируется как прилагательное, образованное от существительного, и существительное в косвенном падеже. В английском языке этот символ получил бы еще дополнительную интерпретацию — существительное в общем (именительном) падеже в атрибутивной функции (a *snake bite*).

Если бы перед исследователем стояла цель сопоставительного исследования языков, то правила интерпретации должны были бы учитывать все способы выражения данного грамматического значения в сопоставляемых языках.

Правила интерпретации зависят от того, составляются ли они для фрагмента грамматики или грамматики языка в целом. Так, если бы перед нами стояла цель общего описания русской грамматики, то нам следовало бы интерпретировать символ R_2V как инфинитив и как отгла-

гольное имя. Однако этому символу мы даем только одну интерпретацию, так как нас интересуют словообразовательные отношения между глаголом и отглагольным именем. Для наших целей также несущественно уточнять морфологическую характеристику основ у слов, соотносимых, в частности, с символами V, N, R_4N, тогда как в общей грамматике необходимо было бы указать на их непроизводность или производность.

Ниже приводятся правила, составленные [11] нами для интерпретации описанной выше модели:

Символ	Интерпретация символа	Сокращенное обозначение
$N_{(1,\ 2)}$	Существительное (местоимение) в именительном падеже в функции подлежащего	$С_{(1,\ 2)}$
V	Глагол в личной форме	Г
A; $R_3(R_4A)$	Непроизводное прилагательное	П
D_2	Непроизводное наречие	Н
R_2V	Отглагольное существительное	$С_г$
$R_1(R_2V)$	Присвязочное отглагольное существительное, употребляющееся со связками типа:	
	1) (*быть*) *учителем*, *служить украшением*	$св_1С_г$
	2) (*быть*) *в отчаянии*, *находиться под угрозой*	$св_2С_г$
	3) *делать поворот*, *оказывать влияние*	$св_3С_г$
	4) (*быть*) *хорошего воспитания*	$св_4С_г$
$R_3N_{(1)}$	Существительное в родительном падеже без предлога	$С^{(1)}_{рп}$
	Остальные косвенные падежи существительного (с предлогами и без предлогов)	$С^{(1)}_{косв}$
	Отыменное прилагательное *	$П^{(1)}_с$
$R_3(R_4N_1)$	Существительное в родительном падеже без предлога	$С^1_{рп}$
	Отыменное прилагательное	$П^1_с$
R_3N_2	Существительное в родительном падеже без предлога	$С^2_{рп}$
	Остальные косвенные падежи существительного (с предлогами и без предлогов)	$С_{косв}$
	Отыменное прилагательное	$П^2_с$
	Именная основа, входящая в состав сложных слов типа *работорговля* и *небожитель*.	$С^2_{осн}$
$R_3(R_4N_2)$	Существительное в родительном падеже без предлога	$С^2_{рп}$
	Отыменное прилагательное	$П^2_с$
	Именная основа, входящая в состав сложных слов типа *работорговля* и *небожитель*	$С^2_{осн}$

* Символ R_3N_1 можно было бы еще интерпретировать и как именную основу, входящую в состав сложных слов типа *солнцестояние* (*Солнце стоит высоко — Высокое солнцестояние*). Однако из-за единичности и непродуктивности таких образований при переходе к реальному языку такая интерпретация породила бы очень большое число неграмматичных предложений.

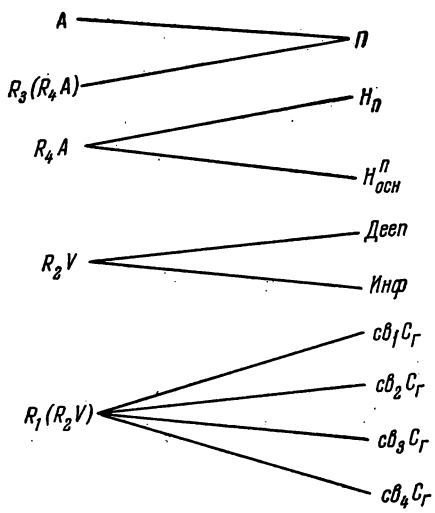
[11] Правила интерпретации составлены нами для простых членов ядерных и деривационных цепочек.

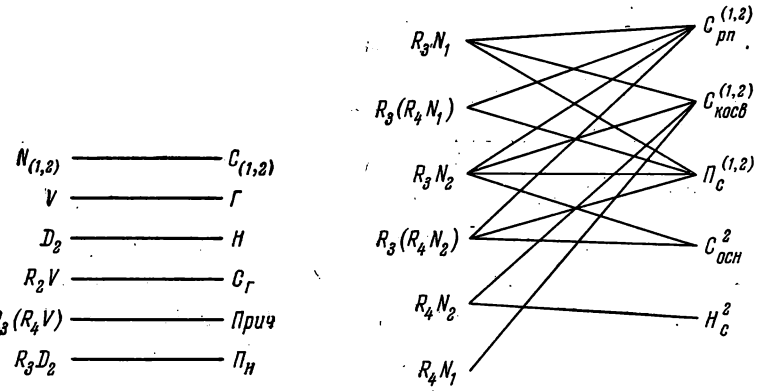
Окончание

Символ	Интерпретация символа	Сокращенное обозначение
$R_3(R_4V)$	Причастие *	Прич
R_3D_2	Прилагательное, образованное от наречия	$П_н$
R_4A	Наречие, образованное от прилагательного	$Н_п$
	Именная основа, входящая в состав сложных слов типа *громкоговоритель, чистописание*	$Н^п_{осн}$
R_4V	Деепричастие	Дееп
	Инфинитив	Инф
R_4N_2	Существительное в косвенном падеже, кроме родительного беспредложного, предложные сочетания с существительными	$С^2_{косв}$
	Наречие, образованное от существительного	$Н^2_с$
R_4N_1	Существительное в косвенном падеже, кроме родительного беспредложного, предложные сочетания с существительными	$С^1_{косв}$

* Если символ $R_3(R_4V)$ интерпретировать еще и как отглагольное прилагательное типа *хриплый, вращательный*, то можно было бы учесть соотношения типа: *Вал движется, вращаясь* → ... *вращательное движение вала... Хрипя запел граммофон* → ... *хриплое граммофонное пение...* Однако при переходе к реальному языку такая интерпретация создает очень большое число неграмматичных предложений.

Из правил интерпретации следует, что между символами модели и грамматическими объектами реального языка устанавливаются три типа соответствия: одно-однозначное, одно-многозначное и много-многозначное. Последнее можно представить в виде следующих графов:





Помимо правил интерпретации для перехода от абстрактной модели к реальному языку необходимо разработать систему других правил, в частности — правила лексического воплощения слов, правила порядка следования слов, морфонематические правила и ряд других.

Однако такие правила рационально разрабатывать для грамматики в целом, а не для ее отдельного фрагмента. Поэтому реальные предложения русского языка, приводимые ниже, должны рассматриваться как результат порождения только с точки зрения их грамматической структуры.

Переформулируем цепочки 1—5 и «а»—«ф» в терминах, установленных нашими правилами интерпретации. Цепочки «а» и «б», «в» и «г», «д» и «е» и т. д. объединим попарно по принципу совпадения структуры доминирующего члена.

1) $С + Г + Н$

$$\left\{\begin{matrix} С_{рп} \\ Р_{косв} \\ П_{С} \end{matrix}\right\} + С_Г + \left\{\begin{matrix} П_Н \\ Н \end{matrix}\right\} \quad \text{(а, б)}$$

$$С + \left\{\begin{matrix} св_1С_Г \\ св_2С_Г \\ св_3С_Г \\ св_4С_Г \end{matrix}\right\} + \left\{\begin{matrix} П_Н \\ Н \end{matrix}\right\} \quad \text{(в, г)}$$

2) $С + Г + Н_П$

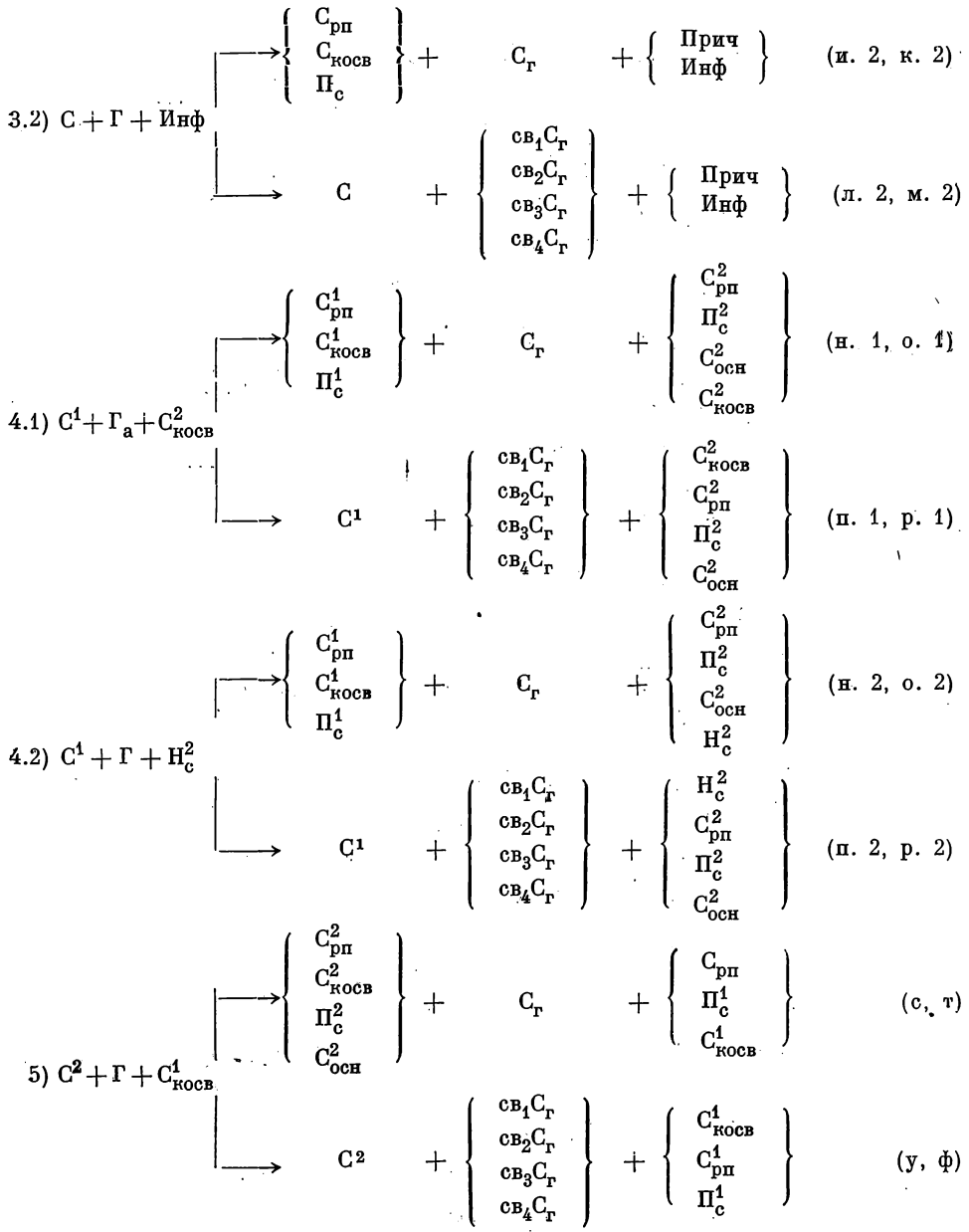
$$\left\{\begin{matrix} С_{рп} \\ С_{косв} \\ П_{С} \end{matrix}\right\} + С_Г + \left\{\begin{matrix} П \\ Н_П \\ Н^{П}_{осн} \end{matrix}\right\} \quad \text{(д, е)}$$

$$С + \left\{\begin{matrix} св_1С_Г \\ св_2С_Г \\ св_3С_Г \\ св_4С_Г \end{matrix}\right\} + \left\{\begin{matrix} П \\ Н_П \\ Н^{П}_{осн} \end{matrix}\right\} \quad \text{(ж, з)}$$

3.1) $С + Г + Дееп$

$$\left\{\begin{matrix} С_{рп} \\ С_{косв} \\ П_{С} \end{matrix}\right\} + С_Г + \left\{\begin{matrix} Прич \\ Дееп \end{matrix}\right\} \quad \text{(и. 1, к. 1)}$$

$$С + \left\{\begin{matrix} св_1С_Г \\ св_2С_Г \\ св_3С_Г \\ св_4С_Г \end{matrix}\right\} + \left\{\begin{matrix} Прич \\ Дееп \end{matrix}\right\} \quad \text{(л. 1, м. 1)}$$



После перекодировки символов оказалось, что цепочки «н.1», «о.1» и «с», «т» дублируют друг друга (порядок расположения символов в формуле на данном уровне описания языка не имеет значения). Для того, чтобы на уровне реального языка различать цепочки 4 → «н.1», «о.1» и 5 → «с», «т», припишем залоговые характеристики символу Г в этих цепочках. В первом случае это будет глагол в активном залоге, обозначим его $Г_а$, во втором — глагол в страдательном залоге, обозначим его $Г_б$.

Раскрыв скобки, развернем данные формулы, исчислив все возможные сочетания элементов одной скобки со всеми элементами другой. Результат такого исчисления считаем списком искомых трансформаций.

Иллюстрируя трансформации, учитываем дополнительное требование, диктуемое данным уровнем описания языка, а именно требование тождественности лексических морфем в исходном и производном трансформе.

Исходный трансформ	Производные трансформы		Иллюстрация
1) $С + Г + Н$	а. 1	$С_{рп} + П_{н} + С_{г}$	*Вчера заседал местком* → *. . . вчерашнее заседание месткома...*
	а. 2	$С_{косв} + П_{н} + С_{г}$	*Здесь пасется скот.* → *. . . здешний выпас для скота...*
	а. 3	$П_{с} + П_{н} + С_{г}$	*Друзья беседовали вчера.* → *... вчерашняя дружеская беседа ...*
	б. 1	$С_{рп} + С_{г} + Н$	*Сюда прибыл самолет.* → *... прибытие сюда самолета ...*
	б. 2	$С_{косв} + С_{г} + Н$	*Здесь раскололось зеркало.* → *... здесь осколки от зеркала ...*
	б. 3	$П_{с} + С_{г} + Н$	*Всюду лаяли собаки.* → *... собачий лай всюду ...*
	в. 1	$С + св_{1}С_{г} + Н$	*Она всегда обманывала.* → *Она всегда была обманщицей.*
	в. 2	$С + св_{2}С_{г} + Н$	*Там пасутся коровы.* → *Там коровы на пастбище.*
	в. 3	$С + св_{3}С_{г} + Н$	*Установка теперь испытывается.* → *Установка теперь проходит испытание*
	в. 4	$С + св_{4}С_{г} + Н$	*
	г. 1	$С + св_{1}С_{г} + П_{н}$	*Он здесь преподает.* → *Он здешний преподаватель*
	г. 2	$С + св_{2}С_{г} + П_{н}$	*Здесь зимовали оленеводы.* → *Оленеводы были на здешнем зимовье*
	г. 3	$С + св_{3}С_{г} + П_{н}$	*Кто докладывал вчера?* → *Кто делал вчерашний доклад?*
	г. 4	$С + св_{4}С_{г} + П_{н}$	*Здесь изготовляются ткани.* → *Ткани — здешнего изготовления*
2) $С + Г + Н_{п}$	д. 1	$С_{рп} + П + С_{г}$	*Тихо шелестят листья.* → *...тихий шелест листьев ...*
	д. 2	$С_{косв} + П + С_{г}$	*Фонарь тускло светит.* → *... тусклый свет от фонаря ...*
	д. 3	$П_{с} + П + С_{г}$	*Уныло шумит море.* → *... унылый морской шум...*
	е. 1	$С_{рп} + С_{г} + Н_{п}$	*Яйца сварены вкрутую.* → *... варка яиц вкрутую...*

Продолжение

Исходный трансформ	Производные трансформы		Иллюстрация
2) $С + Г + Н_п$	е. 2	$С_{косв} + С_г + Н_п$	*Самолет летит вслепую.* → *... полет на самолете вслепую ...*
	е. 3	$П_с + С_г + Н_п$	*Волк завыл вдалеке.* → *... волчий вой вдалеке ...*
	е. 4	$С_{рп} + С_г + Н_п^{осн}$	*Подхалим низко кланяется.* → *... низкопоклонство подхалима ...*
	е. 5	$С_{косв} + С_г + Н_п^{осн}$	*
	е. 6	$П_с + С_г + Н_п^{осн}$	*Праздно шатались пижоны.* → *... пижонское праздношатательство..,*
	ж. 1	$С + св_1С_г + Н_п$	*Он по-прежнему не платит.* → *Он неплательщик по-прежнему*
	ж. 2	$С + св_2С_г + Н_п$	*Книга давно печатается.* → *Книга давно в печати*
	ж. 3	$С + св_3С_г + Н_п$	*Он часто ошибается.* → *Он часто делает ошибки*
	ж. 4	$С + св_4С_г + Н_п$	*
	ж. 5	$С + св_1С_г + Н_п^{осн}$	*Корова отелилась впервые.* → *Корова — первотелка*
	ж. 6	$С + св_2С_г + Н_п^{осн}$	*
	ж. 7	$С + св_3С_г + Н_п^{осн}$	*
	ж. 8	$С + св_4С_г + Н_п^{осн}$	*
	з. 1	$С + св_1С_г + П$	*Он прекрасно работает.* → *Он прекрасный работник*
	з. 2	$С + св_2С_г + П$	*Она глубоко отчаялась.* → *Она в глубоком отчаянии*
	з. 3	$С + св_3С_г + П$	*Он поступил опрометчиво.* → *Он совершил опрометчивый поступок*
	з. 4	$С + св_4С_г + П$	*Пальто красиво покроено.* → *Пальто красивого покроя*
3.1) $С + Г +$ Деепр	и. 1.1	$С_{рп} +$ Прич $+ С_г$	*Девушка глядит, тоскуя.* → *... тоскующий взгляд девушки ...*
	и. 1.2	$С_{косв} +$ Прич $+ С_г$	*Город строится, расширяясь.* → *... расширяющееся строительство в городе ...*
	и. 1.3	$П_с +$ Прич $+ С_г$	*Лиса бежит, петляя.* → *...лисий петляющий бег...*
	к. 1.1	$С_{рп} + С_г +$ Деепр	*

Продолжение

Исходный трансформ	Производные трансформы		Иллюстрация
3.1) С + Г + Деепр	к. 1.2	$С_{косв} + С_г$ + Деепр	*
	к. 1.3	$П_с + С_г$ + Деепр	*
	л. 1.1	С + $св_1С_г$ + Деепр	*
	л. 1.2	С + $св_2С_г$ + Деепр	*
	л. 1.3	С + $св_3С_г$ + Деепр	*Раненый стонал, не переставая.* → *Раненый, не переставая, издавал стоны*
	л. 1.4	С + $св_4С_г$ + Деепр	*
	м. 1.1	С + $св_1С_г$ + Прич	*Она учится, работая.* → *Она работающая ученица*
	м. 1.2	С + $св_2С_г$ + Прич	*
	м. 1.3	С + $св_3С_г$ + Прич	*Он глядит, недоумевая.* → *Он бросает недоумевающие взгляды*
	м. 1.4	С + $св_4С_г$ + Прич	*
3.2) С + Г + Инф	и. 2.1	$С_{рп}$ + Прич + $С_г$	*
	и. 2.2	$С_{косв}$ + Прич + $С_г$	*
	и. 2.3	$П_с$ + Прич + $С_г$	*
	к. 2.1	$С_{рп} + С_г$ + Инф	*Мальчик стремится учиться.* → *...стремление мальчика учиться...*
	к. 2.2	$С_{косв} + С_г$ + Инф	*Они не желают трудиться.* → *Нежелание трудиться для них...*
	к. 2.3	$П_с + С_г$ + Инф	*Человек мечтал летать.* → *...человеческая мечта летать ...*
	л. 2.1	С + $св_1С_г$ + Инф	*Он любит пошутить.* → *Он любитель пошутить*
	л. 2.2	С + $св_2С_г$ + Инф	*Постройка грозит развалиться.* → *Постройка под угрозой развалиться*
	л. 2.3	С + $св_3С_г$ + Инф	*Он намеревался выступить.* → *Он имел намерение выступить*
	л. 2.4	С + $св_4С_г$ + Инф	*
	м. 2.1	С + $св_1С_г$ + Прич	*
	м. 2.2	С + $св_2С_г$ + Прич	*
	м. 2.3	С + $св_3С_г$ + Прич	*
	м. 2.4	С + св $С_г$ + Прич	*

Продолжение

Исходный трансформ	Производные трансформы		Иллюстрация
4.1) $С^1 + Г_а + С^2_{косв}$	н. 1.1	$С^1_{рп} + С^2_{рп} + С_г$	*
	н. 1.2	$С^1_{косв} + С^2_{рп} + С_г$	*Самолет доставил почту.* → *... доставка почты самолетом (на самолете) ...*
	н. 1.3	$П^1_с + С^2_{рп} + С_г$	*Завод выпускает продукцию.* → *... выпуск заводской продукции ...*
	н. 1.4	$С^1_{рп} + П^2_с + С_г$	*Рыба весит килограмм.* → *... килограммовый вес рыбы ...*
	н. 1.5	$С^1_{косв} + П^2_с + С_г$	*Украина поставляет хлеб.* → *... хлебные поставки с Украины ...*
	н. 1.6	$П^1_с + П^2_с + С_г$	*Ташкент производит ковры.* → *...ташкентское ковровое производство...*
	н. 1.7	$С^1_{рп} + С^2_{осн} + С_г$	*Больной помешался умом.* → *... умопомешательство больного ...*
	н. 1.8	$С^1_{косв} + С_{осн} + С_г$	*Успехи кружат голову.* → *... головокружение от успехов ...*
	н. 1.9	$П^1_с + С^2_{осн} + С_г$	*Америка торговала рабами* → *... американская работорговля ...*
	о 1.1	$С^1_{рп} + С_г + С^2_{косв}$	*Люди стремятся к счастью.* → *... стремление людей к счастью ...*
	о. 1.2	$С^1_{косв} + С_г + С^2_{косв}$	*Шкаф поцарапал стену.* → *... царапина от шкафа на стене ...*
	о. 1.3	$П^1_с + С_г + С^2_{косв}$	*Народы борются против колониализма.* → *... народная борьба против колониализма ...*
	п. 1.1	$С^1 + св_1С_г + С^2_{косв}$	*Он борется за мир.* → *Он борец за мир*
	п. 1.2	$С^1 + св_2 + С^2_{косв}$	*Трактора пашут за лесом.* → *Трактора на пашне за лесом*
	п. 1.3	$С^1 + св_3С_г + С^2_{косв}$	*Поэзия влияет на умы.* → *Поэзия оказывает влияние на умы*
	п. 1.4	$С^1 + св_4С_г + С^2_{косв}$	*

Продолжение

Исходный трансформ	Производные трансформы		Иллюстрация
4.1) $С^1 + Г_а + С^2_{косв}$	р. 1.1	$С^1 + св_1С_г + С^2_{рп}$	*Он водит автобус.* → *Он водитель автобуса*
	р. 1.2	$С^1 + св_2С_г + С^2_{рп}$	*Они убирают урожай.* → *Они на уборке урожая*
	р. 1.3	$С^1 + св_3С_г + С^2_{рп}$	*Мы перестраиваем промышленность.* → *Мы осуществляем перестройку промышленности*
	р. 1.4	$С^1 + св_4С_г + С^2_{рп}$	*
	р. 1.5	$С^1 + св_1С_г + П^1_с$	*Он угождает дамам.* → *Он дамский угодник*
	р. 1.6	$С^1 + св_2С_г + П^2_с$	*Он спускается с горы.* → *Он на горном спуске*
	р. 1.7	$С^1 + св_3С_г + П^2_с$	*Мы беседовали два часа.* → *Мы имели двухчасовую беседу*
	р. 1.8	$С^1 + св_4С_г + П^2_с$	*
	р. 1.9	$С^1 + св_1С_г + С^2_{осн}$	*Он дает взятку.* → *Он взяткодатель*
	р. 1.10	$С^1 + св_2С_г + С^2_{осн}$	*Они заготовляют сено.* → *Они на сенозаготовках*
	р. 1.11	$С^1 + св_3С_г + С^2_{осн}$	*Он провел электричество.* → *Он сделал электропроводку*
	р. 1.12	$С^1 + св_4С_г + С^2_{осн}$	*
4.2) $С^1 + Г + Н^2_с$	н. 2.1	$С^1_{рп} + С^2_{рп} + С_г$	*
	н. 2.2	$С^1_{косв} + С^2_{рп} + С_г$	*
	н. 2.3	$П^1_с + С^2_{рп} + С_г$	*
	н. 2.4	$С^1_{рп} + П^2_с + С_г$	*Колеса стучат вдали.* → *...дальний стук колес...*
	н. 2.5	$С^1_{косв} + П^2_с + С_г$	*Судно кренится набок.* → *... боковой крен у судна ...*
	н. 2.6	$П^1_с + П^2_с + С_г$	*Тетка врет безудержу.* → *... безудержное теткино вранье ...*
	н. 2.7	$С^1_{рп} + С^2_{осн} + С_г$	*
	н. 2.8	$С^1_{косв} + С^2_{осн} + С_г$	*
	н. 2.9	$П^1_с + С^2_{осн} + С_г$	*

Продолжение

Исходный трансформ	Производные трансформы		Иллюстрация
4.2) $С^1 + Г + Н^2_с$	о. 2.1	$С^1_{рп} + С_г + Н^2_с$	*Овощи заготовляются впрок.* → *... заготовка овощей впрок ...*
	о. 2.2	$С^1_{косв} + С_г + Н^2_с$	*Арбуз режется сверху.* → *... надрез на арбузе сверху ...*
	о. 2.3	$П^1_с + С_г + Н^2_с$	*Птицы гогочут вдали.* → *... птичий гогот вдали...*
	п. 2.1	$С^1 + св_1С_г + Н^2_с$	*Он по-прежнему прогуливает.* → *Он по-прежнему прогульщик*
	п. 2.2	$С^1 + св_2С_г + Н^2_с$	*Я уезжала летом.* → *Летом я была в отъезде*
	п. 2.3	$С^1 + св_3С_г + Н^2_с$	*Он взглянул исподлобья.* → *Он бросил взгляд исподлобья*
	п. 2.4	$С^1 + св_4С_г + Н^2_с$	*
	р. 2.1	$С^1 + св_1С_г + С^2_{рп}$	*
	р. 2.2	$С^1 + св_2С_г + С^2_{рп}$	*
	р. 2.3	$С^1 + св_3С_г + С^2_{рп}$	*
	р. 2.4	$С^1 + св_4С_г + С^2_{рп}$	*
	р. 2.5	$С^1 + св_1С_г + П^2_с$	*Он пьет беспросыпу.* → *Он беспросыпный пьяница*
	р. 2.6	$С^1 + св_2С_г + П^2_с$	*Они заседали утром.* → *Они были на утреннем заседании*
	р. 2.7	$С^1 + св_3С_г + П^2_с$	*Они сеяли весной.* → *Они проводили весенний сев*
	р. 2.8	$С^1 + св_4С_г + П^2_с$	*Травы собирались летом.* → *Травы — летнего сбора*
	р. 2.9	$С^1 + св_1С_г + С^2_{осн}$	*
	р. 2.10	$С^1 + св_2С_г + С^2_{осн}$	*
	р. 2.11	$С^1 + св_3С_г + С^2_{осн}$	*
	р. 2.12	$С^1 + св_4С_г + С^2_{осн}$	*
5) $С^2 + Г_б + С^1_{косв}$	с. 1	$С^2_{рп} + С^1_{рп} + С_г$	*

Продолжение

Исходный трансформ		Производные трансформы	Иллюстрация
5) $C^2 + Г_б + C^1_{косв}$	с. 2	$C^2_{косв} + C^1_{рп} + C_г$	*
	с. 3	$П^2_с + C^1_{рп} + C_г$	*На фабрике делаются ковры.* → *...ковровые изделия фабрики ...*
	с. 4	$C^2_{осн} + C^1_{рп} + C_г$	*Паровозы строятся в Луганске.* → *... паровозостроение Луганска ...*
	с. 5	$C^2_{рп} + П^1_с + C_г$	*Пшеница поставляется Сибирью.* → *... сибирские поставки пшеницы ...*
	с. 6	$C^2_{косв} + П^1_с + C_г$	*Шкатулка инкрустирована серебром.* → *... серебряная инкрустация на шкатулке ...*
	с. 7	$П^2_с + П^1_с + C_г$	*Трикотаж производится в Косине.* → *Трикотаж косинского производства*
	с. 8	$C^2_{осн} + П^1_с + C_г$	*В колхозе разводятся пчелы.* → *... колхозное пчеловодство ...*
	т. 1	$C^2_{рп} + C_г + C^1_{косв}$	*Больные осматриваются врачом.* → *... осмотр больных врачом ...*
	т. 2	$C^2_{косв} + C_г + C^1_{косв}$	*Елка украшена фольгой.* → *... украшения из фольги на елке ...*
	т. 3	$П^2_с + C_г + C^1_{косв}$	*В селе производятся бочки.* → *бочарное производство на селе ...*
	т. 4	$C^2_{осн} + C_г + C^1_{осн}$	*Звуки записываются на магнитофоне.* → *... звукозапись на магнитофоне ...*
	у. 1	$C^2 + св_1 C_г + C^1_{косв}$	*Он воспитывался в приюте.* → *Он был воспитанником в приюте*
	у. 2	$C^2 + св_2 C_г + C^1_{косв}$	*Дело рассматривалось в суде.* → *Дело было на рассмотрении в суде*
	у. 3	$C^2 + св_г C_г + C^1_{косв}$	*Его поведение осуждалось общественностью.* → *Его поведение получило осуждение со стороны общественности*
	у. 4	$C^2 + св_4 C_г + C^1_{косв}$	*

Окончание

Исходный трансформ	Производные трансформы		Иллюстрация
5) $С^2 + Г_б + С^1_{косв}$	ф. 1	$С^2 + св_1С_г + С^1_{рп}$	*Радио изобретено Поповым.* → *Радио — изобретение Попова*
	ф. 2	$С^2 + св_2С_г + С^1_{рп}$	*Вопрос обсуждался на дирекции.* → *Вопрос был на обсуждении дирекции*
	ф. 3	$С^2 + св_3С_г + С^1_{рп}$	*Обращение было поддержано массами.* → *Обращение получило поддержку масс*
	ф. 4	$С^2 + св_4С_г + С^1_{рп}$	*Трактор выпущен в 1929 году.* → *Трактор выпуска 1929 года*
	ф. 5	$С^2 + св_1С_г + П^1_с$	*Картина приобретена школой.* → *Картина — школьное приобретение*
	ф. 6	$С^2 + св_2С_г + П^1_с$	*Он был сослан царем.* → *Он был в царской ссылке*
	ф. 7	$С^2 + св_3С_г + П^1_с$	*Он был ранен пулей.* → *Он получил пулевое ранение*
	ф. 8	$С^2 + св_4С_г + П^1_с$	*Он воспитан в Петербурге.* → *Он петербургского воспитания*

Оценка результатов исчисления. Полученная модель фрагмента русской грамматики имеет вполне удовлетворительную объяснительную силу. Из 143 трансформаций, моделирующих возможные преобразования, которым подвергается глагольная синтагма при обращении глагола в отглагольное имя, 99 находит интерпретацию в русском языке.

Для 42 трансформаций, отмеченных в списке звездочкой, удовлетворительной интерпретации подобрать не удалось. Это, например, трансформации типа

$$С + Г + Инф \longrightarrow \left\{ \begin{matrix} С_{рп} \\ С_{косв} \\ П_с \end{matrix} \right\} + Прич + С_г$$

и другие, против которых стоят «пустые клетки». Пустая клетка означает невозможность таких преобразований, как, например: *Он стремится учиться* → **Его учащееся стремление*. В данном случае имеется в виду недопустимость таких преобразований с точки зрения грамматических возможностей русского языка. Трансформы такого типа следует считать грамматически не отмеченными.

От грамматической отмеченности следует отличать лексическую, о которой можно говорить при исследовании на трансформируемость конкретных фраз в пределах грамматических возможностей языка.

Подвергнем, например, конкретную фразу *Украина поставляет хлеб* возможным преобразованиям, обобщенным в цепочках «н.1», «о.1», «п.1», «у», «ф»:

н. 1.1	$С^1_{рп} + С^2_{рп} + С_г$	*
н. 1.2	$С^1_{косв} + С^2_{рп} + С_г$	*Поставки хлеба Украиной (с Украины)*
н. 1.3	$П^1_с + С^2_{рп} + С_г$	*Украинские поставки хлеба*
н. 1.4	$С^1_{рп} + П^2_с + С_г$	*Хлебные поставки Украины*
н. 1.5	$С^1_{косв} + П^2_с + G_г$	*Хлебные поставки с Украины*
н. 1.6	$П^1_с + П^2_с + С_г$	*Украинские хлебные поставки*
н. 1.7	$С^1_{рп} + С^2_{осн} + С_г$	*Хлебопоставки Украины*
н. 1.8	$С^1_{косв} + С^2_{осн} + С_г$	*Хлебопоставки с Украины*
н. 1.9	$П^1_с + С^2_{осн} + С_г$	*Украинские хлебопоставки*
о. 1.1	$С^1_{рп} + С_г + С^2_{косв}$	*
о. 1.2	$С^1_{косв} + С_г + С^2_{косв}$	*
о. 1.3	$П^1_с + С_г + С^2_{косв}$	*
п. 1.1	$С^1 + св_1С_г + С^2_{косв}$	*
п. 1.2	$С^1 + св_2С_г + С^2_{косв}$	*
п. 1.3	$С^1 + св_3С_г + С^2_{косв}$	*Украина производит поставки хлебом*
п. 1.4	$С^1 + св_4С_г + С^2_{косв}$	*
р. 1.1	$С^1 + св_1С_г + С^2_{рп}$	*Украина — поставщик хлеба*
р. 1.2	$С^1 + св_2С_г + С^2_{рп}$	*
р. 1.3	$С^1 + св_3С_г + С^2_{рп}$	*Украина производит поставку хлеба*
р. 1.4	$С^1 + св_4С_г + С^2_{рп}$	*
р. 1.5	$С^1 + св_1С_г + П^2_с$	*Украина — хлебный поставщик*
р. 1.6	$С^1 + св_2С_г + П^2_с$	*
р. 1.7	$С^1 + св_3С_г + П^2_с$	*Украина производит хлебные поставки*
р. 1.8	$С^1 + св_4С_г + П^2_с$	*
р. 1.9	$С^1 + св_1С_г + С^2_{осн}$	*Украина является хлебопоставщиком*
р. 1.10	$С^1 + св_2С_г + С^2_{осн}$	*
р. 1.11	$С^1 + св_3С_г + С^2_{осн}$	*Украина производит хлебопоставки*
р. 1.12	$С^1 + св_4С_г + С^2_{осн}$	*
у. 1	$С^2 + св_1С_г + С^2_{косв}$	*Хлеб — (это) поставки с Украины*
у. 2	$С^2 + св_2С_г + С^1_{косв}$	*Хлеб — из поставок с Украины*
у. 3	$С^2 + св_3С_г + С^1_{косв}$	*
у. 4	$С^2 + св_4С_г + С^1_{косв}$	*
ф. 1	$С^2 + св_1С_г + С^1_{рп}$	*Хлеб — (это) поставки Украины*
ф. 2	$С^2 + св_2С_г + С^1_{рп}$	*
ф. 3	$С^2 + св_3С_г + С^1_{рп}$	*
ф. 4	$С^2 + св_4С_г + С^1_{рп}$	*
ф. 5	$С^2 + св_1С_г + П^1_с$	*Хлеб — (это) украинские поставки*
ф. 6	$С^2 + св_2С_г + П^1_с$	*Хлеб — из украинских поставок*
ф. 7	$С^2 + св_3С_г + П^1_с$	*
ф. 8	$С^2 + св_4С_г + П^1_с$	*Хлеб — украинской поставки*

Пустые клетки в данном случае означают лексическую невозможность такой трансформации именно для данной фразы. Можно сказать, что предложения типа **Поставки хлебом Украины* являются лексически не отмеченными.

Обладая вполне приличной объяснительной силой (70%), полученная система трансформаций, моделирующая связи глагола с отлагательным именем на синтаксическом уровне, имеет в то же время ряд существенных недостатков.

В частности, принятая в правилах интерпретации трехпадежная система (именительный, родительный беспредложный, косвенный) является довольно грубым приближением к фактам реального языка. Такая система не позволяет отличить случаи типа *Я пишу брату* → *. . . мое письмо брату. . .*, где сохраняется один и тот же падеж управляемого существительного, от случаев типа *Игрушки украшают елку* → *. . . игрушечные украшения на елке. . .*, где косвенный падеж существительного меняется (на любой, кроме родительного беспредложного). В обоих случаях управляемый член записывается как $С^2_{косв}$ в исходном и производном трансформе (см. цепочки «о» и «п»). Кроме того, получившаяся система трансформаций не отражает случаев типа *Он жаждет справедливости* → *. . . его жажда справедливости. . .*, т. е. такие преобразования, при которых сохраняется родительный беспредложный падеж управляемого существительного.

Оба отмеченных недостатка объясняются интерпретацией символа R_4N как любого косвенного падежа, кроме родительного беспредложного (значение которого приписывается символам R_3N и $R_3(R_4N)$.

Вместе с тем устранение этих недостатков, связанных с трудностью разработки адекватных правил интерпретации, нерационально усложнило бы всю систему.

Некоторые выводы для семантики. Вернемся к формулам на стр. 250—251, получившимся после переформулировки цепочек 1—5 и «а»—«ф» в терминах, установленных нашими правилами интерпретации.

Перепишем эти формулы, перегруппировав их по принципу структурного типа доминирующего члена, и получим таким образом пять групп формул. Здесь достаточно привести лишь первую и вторую группы (для типов $С_г$ и $св_1С_г$); остальные три группы (для типов $св_2С_г$, $св_3С_г$ и $св_4С_г$) только тем отличаются от второй, что на месте члена $св_1С_г$ соответственно стоят члены $св_2С_г$, $св_3С_г$ и $св_4С_г$.

Группа 1

$$С + Г + Н \longrightarrow \left\{ \begin{matrix} С_{рп} \\ С_{косв} \\ П_с \end{matrix} \right\} + С_г + \left\{ \begin{matrix} П_н \\ \\ Н \end{matrix} \right\}$$

$$С + Г + Н_п \longrightarrow \left\{ \begin{matrix} С_{рп} \\ С_{косв} \\ П_с \end{matrix} \right\} + С_г + \left\{ \begin{matrix} П \\ Н_п \\ Н^п_{осн} \end{matrix} \right\}$$

$$С + Г + Дееп \longrightarrow \left\{ \begin{matrix} С_{рп} \\ С_{косв} \\ П_с \end{matrix} \right\} + С_г + \left\{ \begin{matrix} Дееп \\ \\ Прич \end{matrix} \right\}$$

$$С + Г + Инф \longrightarrow \left\{ \begin{matrix} С_{рп} \\ С_{косв} \\ П_с \end{matrix} \right\} + С_г + \left\{ \begin{matrix} Инф \\ \\ Прич \end{matrix} \right\}.$$

$$С^1 + Г + С^2_{косв} \longrightarrow \left\{ \begin{array}{l} С^1_{рп} \\ С^1_{косв} \\ П_с \end{array} \right\} + С_г + \left\{ \begin{array}{l} С^2_{рп} \\ С^2_{косв} \\ П^2_с \\ С^2_{осн} \end{array} \right\}$$

$$С^2 + Г + С^1_{косв} \longrightarrow \left\{ \begin{array}{l} С^2_{рп} \\ С^2_{косв} \\ П^2_с \\ С^2_{осн} \end{array} \right\} + С_г + \left\{ \begin{array}{l} С^1_{рп} \\ С^1_{косв} \\ П^1_с \end{array} \right\}$$

Группа 2

$$С + Г + Н \longrightarrow С + св_1С_г + \left\{ \begin{array}{l} Н \\ П_н \end{array} \right\}$$

$$С + Г + Н_п \longrightarrow С + св_1С_г + \left\{ \begin{array}{l} Н_п \\ Н^п_{осн} \\ П \end{array} \right\}$$

$$С + Г + Дееп \longrightarrow С + св_1С_г + \left\{ \begin{array}{l} Дееп \\ Прич \end{array} \right\}$$

$$С + Г + Инф \longrightarrow С + св_1С_г + \left\{ \begin{array}{l} Инф \\ Прич \end{array} \right\}$$

$$С^1 + Г + С^2_{косв} \longrightarrow С^1 + св_1С_г + \left\{ \begin{array}{l} С^2_{рп} \\ С^2_{косв} \\ П^2_с \\ С^2_{осн} \end{array} \right\}$$

$$С^2 + Г + С^2_{косв} \longrightarrow С^2 + св_1С_г + \left\{ \begin{array}{l} С^1_{рп} \\ С^1_{косв} \\ П^1_с \end{array} \right\}$$

Обратим внимание на тот факт, что в каждой группе формул имеются как постоянные, так и переменные члены. Упростим наши формулы, обозначив переменные члены левых трансформов через x, а правых — через y. Тогда вместо шести левых трансформов получим один, а вместо тридцати правых — пять:

$$С^{(1,2)} + Г + x \begin{cases} \longrightarrow (1) & \left\{ \begin{array}{l} С^{(1,2)}_{рп} \\ С^{(1,2)}_{косв} \\ П^{(1,2)}_с \end{array} \right\} + С_г + y \\ \longrightarrow (2) & С^{(1,2)} + св_1С_г + y \\ \longrightarrow (3) & С^{(1,2)} + св_2С_г + y \\ \longrightarrow (4) & С^{(1,2)} + св_3С_г + y \\ \longrightarrow (5) & С^{(1,2)} + св_4С_г + y, \end{cases}$$

где x принимает значения: $x_1 = $ Н; $x_2 = $ $Н_п$; $x_3 = $ Деепр; $x_4 = $ Инф; $x_5 = $ $= С^1_{косв}$; $x_6 = С^2_{косв}$; а y принимает значения:

$$y_1 = \left\{ \begin{array}{c} Н \\ П_н \end{array} \right\}; \quad y_2 = \left\{ \begin{array}{c} П \\ Н_п \\ Н^п_{осн} \end{array} \right\}; \quad y_3 = \left\{ \begin{array}{c} Деепр \\ Прич \end{array} \right\}; \quad y_4 = \left\{ \begin{array}{c} Инф \\ Прич \end{array} \right\};$$

$$y_5 = \left\{ \begin{array}{c} С^2_{рп} \\ С^2_{косв} \\ П^2_с \\ С^2_{осн} \end{array} \right\}; \quad y_6 = \left\{ \begin{array}{c} С^1_{рп} \\ С^1_{косв} \\ П^1_с \end{array} \right\}.$$

При этом x_1 в левой части формулы соответствует y_1 в правой, x_2 в левой — y_2 в правой и т. д. Индексы при постоянных членах $С^{(1,2)}$, $С^{(1,2)}_{рп}$, $С^{(1,2)}_{косв}$ и $П^{(1,2)}_с$ зависят от конкретного значения x. При $x = x_1 \ldots x_4$ индекс равен нулю. При $x = x_5$ С и другие члены выступают с индексом 1 ($С^1$, $С^1_{рп}$ и т. д.). При $x = x_6$ они выступают с индексом 2 ($С^2$, $С^2_{рп}$ и т. д.).

Пять правых трансформов можно было бы свести всего к двум, тогда вся система полученных трансформаций записывалась бы так:

$$С^{(1,2)} + Г + x \left\{ \begin{array}{ll} \longrightarrow (I) & y_6 + С_г + x \\ \longrightarrow (II) & С^{(1,2)} + \left\{ \begin{array}{c} св_1 \\ св_2 \\ св_3 \\ св_4 \end{array} \right\} С_г + y \end{array} \right.$$

Однако для наших целей удобней остановиться на пяти предыдущих формулах. В таком виде глаголу (Г) и отглагольному существительному ($С_г$) в каждой из пяти формул представляется возможным приписать определенные словообразовательные значения:

Формула 1: действие
- → а) опредмеченное действие
- → б) результат действия

Например: а) *Тихо шелестят листья.* → *...тихий шелест листьев...*

б) *Шкаф поцарапал стену.* → *...царапина от шкафа на стене...*

Формула 2: действие
- → а) действующее лицо или предмет
- → б) объект действия

Например: а) *Он любит пошутить.* → *Он любитель пошутить.*

б) *Он у нас воспитывался.* → *Он был нашим воспитанником.*

Значение «действие → объект действия» соответствует варианту

$$С^2 + Г_б + С^1_{косв} \longrightarrow С^2 + св_1 С_г + \left\{ \begin{array}{c} С^1_{рп} \\ С^1_{косв} \\ П^1_с \end{array} \right\} \text{(формула 2)}$$

Остальные варианты соответствуют значению «действие — действующее лицо или предмет».

Формула 3: действие → а) состояние; действие → б) место действия

Например: а) *Книга давно печатается.* → *Книга давно в печати.*
б) *Там пасутся коровы.* → *Там коровы на пастбище.*
Случаи «а» отличаются от случаев «б» фиксированностью предлога, стоящего перед отглагольным существительным (*в печати, под угрозой, в отчаянии, на излечении, на сенозаготовках* и т. д.).

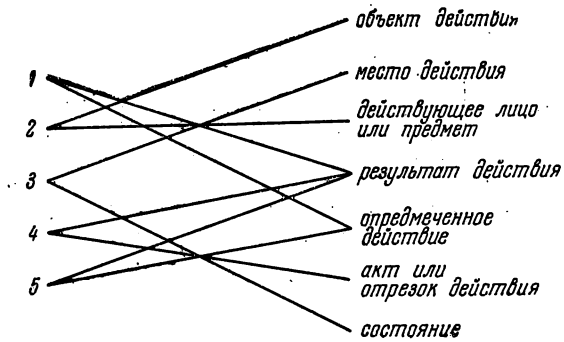
Формула 4: действие → а) акт действия, отрезок действия; действие → б) результат действия

Например: а) *Он глядит, недоумевая.* → *Он бросает недоумевающие взгляды. Мы беседовали два часа.* → *Мы имели двухчасовую беседу.*
б) *Он провел электричество.* → *Он сделал электропроводку.*

Формула 5: действие → а) опредмеченное действие; действие → б) результат действия

Например: а) *Трикотаж производится в Косине.* → *Трикотаж косинского производства.*
б) *Пальто красиво покроено.* → *Пальто — красивого покроя.*

Отношения между типами трансформаций (сведенных к пяти обобщенным формулам) и типами значений отглагольных существительных [12] можно представить в виде графа.



Следующим этапом на пути приближения полученной трансформационной модели к реальному языку может явиться распределение конкретных словообразовательных средств русского языка между всеми вариантами пяти трансформационных формул.

---

[12] Подробнее о связи словообразовательных значений с трансформационными формулами см. П. А. Соболева. О трансформационном анализе словообразовательных отношений. «Трансформационный метод в структурной лингвистике» (в печати).

Автор пользуется случаем, чтобы выразить благодарность товарищам Ю. Д. Апресяну, К. И. Бабицкому, Ю. А. Крутикову, Б. В. Сухотину и С. К. Шаумяну за ценные критические замечания и советы, которые помогли автору при выполнении данной работы.

EXERPIMENT IN TRANSFORMATION CALCULUS ON THE BASIS OF SAUMJAN'S THEORY OF WORD CLASS GENERATION IN THE PROCESS OF GRAMMAR GENERATION

Summary

The paper describes a language generating model involving four operations: substitution, application, rotational transformation and chain transformation.

These operations are applied to an established number of initial strings consisting of nuclear classes the relations between which are those of domination.

Rotational transformation is defined as a transformation resulting from interdependent operations on the dominated classes of a string wherein the original domination is reversed.

Chain transformation is defined as a transformation resulting from an operation on the dominating class of a string wherein the original domination is retained.

The transformation calculus in this paper is based on chain transformation. It is tested for verbal syntagmas the dominating class of which (verb) is changed into a verbal noun. The resultant transformations are interpreted for the Russian language.

А. Б. Долгопольский

КАТЕГОРИЯ ВИДА В РУССКОМ ЯЗЫКЕ И ВЕРОЯТНОСТНЫЙ ХАРАКТЕР СВЯЗИ ОЗНАЧАЕМОГО С ОЗНАЧАЮЩИМ

Настоящая работа явилась результатом исследования семантики видов в русском языке, проводившегося первоначально посредством так называемого трансформационно-переводного метода. Сущность этого метода (в применении к лексике) изложена в моей статье «Изучение лексики трансформационно-переводным методом» [1].

Метод этот сводился к следующему. Каждый язык разбивает семантическое пространство на некоторые участки, из которых каждый есть значение разных языковых знаков (слов, грамматических категорий и др.). Накладываем на одно и то же семантическое пространство разбиение, проводимое каждым языком (из некоторого большого списка языков), а полученные произведением таких разбиений участки принимаем за элементарные смыслы («семы»). Значение знака (морфемы, слова, идиомы, грамматической категории и др.) в каждом языке может быть изображено посредством этих сем, составляющих метаязык семантики.

Однако в ходе исследования, проводившегося на материале видов в современном русском языке, мы пришли к выводу, что такой метод описания семантики для многих элементов языка неприменим. Причина этого — в вероятностном, а не однозначно детерминированном характере связи означаемого с означающим.

Под однозначно детерминированной связью означаемого с означающим мы понимаем следующее: если задано означаемое и даны полные сведения о контексте и внеязыковой ситуации, то мы можем однозначно ответить, какое означающее должно быть в данном языке. Смысл франц. *huit*, англ. *eight* и нем. *acht* в контексте *J'ai huit frères*, *I have eight brothers* и *Ich habe acht Brüder* передается русским *восемь* и никак иначе. Если перед нами в качестве контекста выступает фраза из «Дон Кихота» Сервантеса (т. II, гл. XXI) . . . *acudió con ambas m a n o s a tentar la herida* (оригинал); . . . *eilte . . . herbei, um mit beiden H ä n d e n die Wunde abzutasten* (немецкий перевод Р. Шахта); . . . *came with both h a n d s to feel the wound* (английский перевод П. Мотте под ред. Дж. Г. Локхарта); *lui porta les deux m a i n s sur sa plaie* (французский перевод 1700 г.), то здесь исп. *manos*, нем. *Hände*, англ. *hands* и франц. *mains* непременно передаются русским *руки*: . . . *подбежал, чтобы обеими р у к а м и ощупать рану Басилио* (перевод М. Ватсон), . . . *подбежал к нему, обеими р у к а м и стал ощупывать рану* (перевод под ред. Б. Кржевского и А. Смир-

[1] «Лексикографический сборник», № 5, 1962.

нова), . . . *бросился к нему и, пощупав обеими р у к а м и рану*,. . . (перевод Н. Любимова). В подобных случаях выбор означающего определяется переводными соответствиями и контекстом однозначно.

Не везде, однако, положение столь просто. В ряде случаев (видимо, в большинстве случаев) одно и то же означаемое (один и тот же смысл) может быть в одном и том же контексте передано несколькими разными означающими. При этом в принципе можно найти вероятность употребления любого из этих означающих.

Рассмотрим фразу из «Дон Кихота» (т. II, гл. I). В испанском оригинале:

— *Y a vuesa merced, ¿quién le fía, señor Cura? — dijo don Quijote.*

— *Mi profesión — r e s p o n d i ó el Cura —, que es de guardar secreto.*

В английском переводе (Мотте — Локхарт):

«*And who shall answer for you, Master Curate*?» *answered Don Quixote.*

«*My profession*», *r e p l i e d the curate*, «*which binds me to secrecy*».

В немецком переводе (Р. Шахт):

«*Und wer verbürgt sich für Euer Gnaden, Herr Pfarrer*?» *fragte Don Quijote.*

«*Ich*», *a n t w o r t e t e der Pfarrer*, «*bewahre Geheimnisse von Amts wegen*).

Во французском переводе 1700 г.:

Et qui m'assurera de vous, Monsieur le Curé, dit Don Quichotte.

Mon caractère, r é p o n d i t le curé, qui m'engage à garder le secret à tout le monde.

Здесь испанскому времени pretérito indefinido (*respondió*), английскому Past Indefinite (*replied*), немецкому претериту (*antwortete*) и французскому passé simple (*répondit*) в русском языке может соответствовать как прошедшее время совершенного вида (*ответил*), так и прошедшее несовершенного вида (*отвечал*). Действительно, в переводе под редакцией Б. А. Кржевского и А. А. Смирнова находим форму *ответил*:

— *А за вашу милость, сеньор священник, кто ручается? — спросил Дон Кихот.*

— *Мой сан, — о т в е т и л священник, — повелевающий мне хранить тайны.*

Ту же форму находим в переводе М. Ватсон:

— *А кто поручится мне за вас, сеньор священник? — спросил Дон-Кихот.*

— *Мой сан, — о т в е т и л священник, — обязывающий меня хранить тайны.*

Однако в переводе Н. Любимова употреблена форма *отвечал*:

—*А за вашу милость, сеньор священник, кто поручится? — осведомился Дон Кихот.*

— *Мой сан, повелевающий мне хранить тайны, — о т в е ч а л священник.*

Так же и в переводе В. Карелина:

— *Мое звание, — о т в е ч а л священник, — обязывающее меня хранить доверенные мне тайны.*

Из четырех рассмотренных русских переводов в двух оказалась форма *ответил*, в двух — форма *отвечал*.

Путем постановки соответствующего эксперимента можно получить статистические данные относительно вероятности перевода *ответил* и перевода *отвечал*. Можно было бы предложить значительному числу носителей русского языка (владеющих соответствующей его стилистической разновидностью) перевести данное предложение или заполнить пропуски в уже переведенном предложении (ознакомившись с иностранными соответствиями). При этом, разумеется, этим информантам следует сообщить полный контекст.

Известен случай, когда подобный эксперимент проводился на практике. Кафедре перевода Московского института иностранных языков было поручено провести судебную экспертизу по делу о плагиате, возбужденному С. Серпинским, впервые переведшим роман австралийской писательницы К. С. Причард «Золотые мили», против переводчицы Т. Озерской, опубликовавшей второй перевод того же романа. Один из аргументов истца заключался в следующем: фразу *the fresh air had revived most of the men, and the thought of beer at the nearest pub stimulated sluggish pulses* Т. Озерская, подобно С. Серпинскому, перевела с использованием слов *о кружке пива*. Следовало установить, является ли повторение этого словосочетания плагиатом или оно продиктовано Т. Озерской закономерностями языка. Проводивший экспертизу доцент Я. И. Рецкер поставил следующий опыт: 18 студентам был задан для перевода отрывок из романа «Золотые мили», содержащий указанную фразу. Из 18 студентов 17 перевели *of beer* словами *о кружке пива*. Так была доказана необоснованность обвинения в плагиате (в этом пункте).

В сущности, этот опыт показал, что означаемому, переданному английским словом *beer*, в данном контексте по-русски соответствует означающее *кружка пива* (с вероятностью более 90%) и *пиво* (менее 10%).

Вероятностный характер связи означаемого с означающим делает невозможным повсеместное применение той техники трансформационно-переводного анализа, который предложен в работе «Изучение лексики трансформационно-переводным методом». Границы между зонами разных означающих оказываются не четко определенными, а размытыми, и посредством накладывания таких границ на семантическое пространство нельзя разбить его на конечное число сем.

Изменения, которые в этой связи требуется внести в трансформационно-переводной метод, сводятся к следующему.

1. Трансформационно-переводной анализ в его первом варианте основывался на сличении текста на одном языке с его переводами на ряд других языков. Теперь придется вместо одного текста на каждом языке иметь для каждого языка много однозначных текстов (например, много переводов «Дон Кихота» на русский язык, много его переводов на немецкий язык и т. д., а также много текстов на испанском языке, представляющих собой как бы «перевод» «Дон Кихота», т. е. идентичных «Дон Кихоту» по содержанию).

2. Другое отличие касается результатов исследования. Трансформационно-переводной метод (в первом варианте) должен был привести к построению конечного множества сем, которое служило бы метаязыком для описания семантики любого языка. Теперь обнаруживается, что в связи с размытостью границ между семантическими зонами разных означающих мы не можем добиться такого квантования смысла. Однако в принципе можно предложить процедуру построения приблизительных сем, с заданной точностью сводящих непрерывную шкалу вероятностей к ряду дискретных значений вероятности. Скажем, будет построена сема, которой в английском языке соответствует *blue*, а в русском — на 75% *голубой*, а с вероятностью 25% — *синий*.

Понятно, что такая система сем будет гораздо более громоздка, нежели система сем, которую предполагалось построить трансформационно-переводным методом в первом его варианте.

Поскольку не существует таких текстов, которые имели бы по многу переводов на много языков (да еще и переводы на язык оригинала), эту процедуру целесообразно было бы дополнить экспериментами типа описанного выше. Эти эксперименты пришлось бы проводить с носителями всех привлекаемых к рассмотрению языков. Как видим, процедура

усложняется настолько, что ее выполнение становится практически почти немыслимым (в особенности, если говорить об описании семантики языка в целом).

Поэтому мы вынуждены отказаться от задачи построения метаязыка трансформационно-переводным методом, полагая, что удастся найти другие, менее трудоемкие методы построения такого метаязыка. Возможно, что метаязык удастся построить в результате синтеза логических языков для абстрактных элементов языка [2] с языками, построенными на базе внелингвистической классификации денотатов [3].

Тем не менее, не дожидаясь построения такого метаязыка, мы можем применить идеи трансформационно-переводного метода для типологического описания семантики о д н о г о языка в сопоставлении с другими языками (в частности, и для практических целей машинного перевода и преподавания языков).

Такого рода типологическое описание можно представить в виде алгоритма, позволяющего при заданном означаемом и при заданном окружении определить означающее в описываемом языке (в случаях, когда допускается несколько означающих, разница между которыми в данном окружении иррелевантна, алгоритм должен указать вероятности всех этих означающих).

Определение означающего производится таким путем. Задается весь текст на описываемом (в данном случае на русском) языке, за исключением интересующего нас означающего (в данном случае граммемы вида): *Я никогда не* (*читать* $\vee$ *прочесть*, прош. вр., ед. ч.) *таких книг*; *Когда Павлу, сыну его, было четырнадцать лет — Власову* (*хотеться* $\vee$ *захотеться*, прош. вр., ср. р., ед. ч.) *оттаскать его за волосы.* При этом задается прагматическая характеристика текста (жанр, время написания, время действия и др.).

Чтобы сообщить означаемое, задаются переводы текста на другие языки. На нынешнем этапе исследования в некоторых случаях приходится задавать также дополнительную информацию о ситуации (не содержащуюся в переводах). Приведем пример из области лексики. Если изучаются русские прилагательные цвета и задана фраза *У нее было . . . платье* с переводами *She had a blue dress*, *Tenía un traje azul* и т. д., мы без дополнительной информации (каковой была бы, например, цветная фотография платья) не можем определить, будет ли в этом высказывании по-русски содержаться слово *синее* или *голубое*, а если вероятны оба перевода (если платье не очень темное и не очень светлое), мы не можем без дополнительной информации указать вероятность слова *синее* и слова *голубое*.

Условие наличия дополнительной информации, нарушающее общий принцип описания семантики объективным методом путем обращения только к тексту оригинала и к текстам переводов, возможно, будет снято при дальнейшем совершенствовании методов описания семантики.

Целиком такой алгоритм описания семантики вида (иными словами — употребления вида) русского языка пока не построен, и полное построение его могло бы быть выполнено, видимо, лишь в результате многолетней работы большого коллектива. В настоящем исследовании мы постараемся частично построить алгоритм определения вида по переводам и окружению лишь для некоторого класса высказываний (авторская речь при прямой речи) и укажем общую схему такого алгоритма для остальных классов высказываний.

---

[2] См. стр. 122—124 настоящего сборника (статья Ю. Д. Апресяна).
[3] См. стр. 120—122 той же статьи.

Таблица 1

А. Случаи, когда употребление вида полностью обусловлено окружением

Номера строк	Условия						Ответ
	контекст				означаемое		означающее
	окружение внутри предложения	окружение внутри словоформы	окружение вне предложения	прагматический контекст	переводы на другие языки	дополнительная информация	вид в русском языке
1	*никогда не*	прошедшее время					несов.
2	*всегда*	—					несов.
3	*долго*	нет префикса *про-*					несов.
4	*часто*						несов.
5		настоящее время					несов.
6		глагол *быть*					несов.
7	*слуша-. . ., как* + глагол	прош. вр.					несов.
8	*все еще*						несов.
9	глагол принадлежит авторской речи при прямой речи	глагол *продолжать/продолжить*					несов.

Б. Случаи, когда окружение допускает употребление обоих видов глагола

Номера строк	Условия						Ответ
	контекст				означаемое		означающее
	окружение внутри предложения	окружение внутри словоформы	окружение вне предложения	прагматический контекст	переводы на другие языки	дополнительная информация	вид в русском языке
1		прошедшее время			имперфект в любом из романских языков ∨ имперфект в греческом ∨ прошедшее несовершенное в урду (*-tā* + *-thā* или *-tā* без вспомогательного глагола) ∨ конструкция «основа + *rahā* + *thā*» в урду ∨ англ. Past Continuous		несов.
2	глагол в авторской речи при прямой речи	глагол *продолжить/продолжать* в прош. вр.					несов.
3	То же	глагол *ответить/отвечать* в прош. вр.		современный перевод западного автора XV—XVIII вв.	аорист в греческом, романских языках и в урду (т. е. passé simple во франц., исп. pretérito indefinido, ит. passato remoto, рум. perfect simplu, новогреч. аорист, простое прош. вр. в урду)		несов. с вероятностью 70% и выше (Любимов, Франковский), 10—20% (Горнфельд, Вольпин), около 0% (Кржевский), остальные тексты — сов.

Таблица 1 *(продолжение)*

Номера строк	Условия						Ответ
	контекст				означаемое		означающее
	окружение внутри предложения	окружение внутри словоформы	окружение вне предложения	прагматический контекст	переводы на другие языки	дополнительная информация	вид в русском языке
4	глагол в авторской речи при прямой речи	глаголы *сказать/говорить, возразить/возражать, спросить/спрашивать, начать/начинать, повторять/-ить, воскликать/кнуть* и т. д.		современный перевод западного автора XV—XVIII вв.	аорист в греческом, романских языках и в урду (т. е. passé simple во франц., исп. pretérito indefinido, ит. passato remoto, рум. perfect simplu, новогреч. аорист, простое прош. вр. в урду)		сов.
5	То же			художественная проза XX в. (современная тематика)	То же	большинство информантов (носителей русского языка) утвердительно отвечают на вопрос: «Привел ли автор все слова героя?» — или отвечают отрицательно на вопрос «Повторил ли герой приведенные слова несколько раз?»	сов.
6	»	те же глаголы, что в строке 4		То же	»	большинство информантов дает ответы, противоположные приведенным в строке 5	несов.
7	»	глагол *ответить/отвечать*		»	»	То же	сов. (95—100%), несов. вид (0—5%)

Таблица 1 *(окончание)*

Номера строк	Условия						Ответ	
	контекст				означаемое		означающее	
	окружение внутри предложения	окружение внутри словоформы	окружение вне предложения	прагматический контекст	переводы на другие языки	дополнительная информация	вид в русском языке	
8	глагол в авторской речи при прямой речи	те же глаголы, что в строке 4		художественная проза XX в. (современная тематика)	аорист в греческом, романских языках и в урду (т. е. passé simple во франц., исп. pretérito indefinido, ит. passato remoto, рум. perfect simplu, новогреч. аорист, простое прош. вр. в урду)	отрицательные и утвердительные ответы на вопросы строки 5 распределяются приблизительно поровну, либо большинство информантов допускает обе возможности	сов. (более вероятен), несов. (менее вероятен)	
9	То же	глагол *(по)думать*			То же		несов. не менее вероятен, чем сов.	
10					англ. Past Continuous ∨ плюсквамперфект		несов. и сов.	

СХЕМА АЛГОРИТМА ОПРЕДЕЛЕНИЯ ВИДА ПО ОКРУЖЕНИЮ И ПЕРЕВОДАМ

Общая схема определения вида в русском языке по контексту и переводам высказываний на другие языки может быть представлена в виде таблицы (табл. 1).

Таблица состоит из семи граф.

В графах 1—6 содержатся условия, в последней графе — ответ («совершенный вид», «несовершенный вид» или приближенное указание вероятности совершенного и несовершенного видов в случаях, когда они оба допустимы, а различие их иррелевантно для смысла высказывания).

В графе 1 указывается окружение внутри предложения (грамматические конструкции, наречия, относящиеся к данному глаголу и пр.).

В графе 2 указывается окружение внутри слова (т. е. все, что есть в словоформе: лексема, а также время и иные грамматические категории, кроме, разумеется, вида).

В графе 3 — окружение вне предложения. Пока эта графа пуста, но в будущем, возможно, мы научимся формально описывать контекст и вне предложения.

Графа 4 содержит информацию о прагматическом характере текста («научная работа по математике», «метеорологическая сводка», «современная художественная проза не юмористического характера», «современный перевод западноевропейской прозы XV—XVIII вв.» и пр.).

Графы 5 и 6 имеют общий заголовок «означаемое». Графа 5 указывает некоторые элементы высказываний на других языках, переводящих данное высказывание (в данном случае указываются времена глагола на других языках). Графа 6 содержит дополнительную информацию о ситуации, которая добывается путем опроса информантов. Например, анализируя прилагательные цвета (*She had a blue dress=У нее было . . . платье*), мы задаем информанту вопросы такого рода: «Было ли платье темнее вот этого куска ткани?» (при этом демонстрируется ткань определенного эталона цвета).

В тех случаях, где графы 1—4 (контекст) достаточны для установления вида, графы 5 и 6 (означаемое) не заполняются. Это значит, что при окружениях, отмеченных в графах 1—4, в этих случаях коммутация невозможна, т. е. противопоставление видов отсутствует.

Рассмотрим несколько строк такого алгоритма. Ограничимся двумя прагматическими (стилистическими) разновидностями русского языка: 1) язык оригинальной художественной прозы XX в. (тематика относится к той же эпохе, что и создание произведения), 2) язык современных (XX в.) переводов западной художественной прозы XV—XVIII вв.

А. Случаи, когда употребление вида полностью обусловлено окружением

Строки 1, 2, 4—6 понятны и без примеров. Приведем примеры на остальные строки. Строка 3: *он долго стоял, он долго будет стоять* (но с префиксом *про-*: *он еще долго простоит*). Строка 7: *и слушал, как он . . . бормотал* (А. П. Чехов. Каштанка). Строка 8: *Мойсей Мойсеевич все еще бормотал вполголоса и вздыхал* (А. П. Чехов, Степь, гл. III). О строке 9 см. ниже в разделе Б (строка 2 в табл. 2).

Приведенные здесь строки алгоритма — это всего лишь несколько примеров из весьма обширного перечня тех конкретных условий, которые однозначно определяют выбор вида. Полное описание таких условий представляет значительные трудности. В частности, требуют анализа сложные проблемы зависимости вида глагола от вида других глаголов, сочиненных с ним в качестве однородных членов. Эта зависимость, по-видимому, не одинакова при разном лексическом составе однородных глаголов: можно сказать *он улыбнулся и продолжал*, но нельзя сказать *он улыбнулся и шел*. Не исследована также зависимость вида от вида глагола предложения, сочиненного с данным, и от вида глагола соседнего предложения.

Б. Случаи, когда окружение допускает употребление обоих видов глагола

Среди этих случаев мы рассмотрим лишь одну частную область употребления — вид глагола авторской речи при прямой речи.

Вот как выглядят соответствующие строки алгоритма определения означающего по означаемому и контексту:

Примеры и комментарии

Строка 1. Попробуем определить, какой вид глагола должен быть в следующей фразе:

— *Чтобы понять, почему люди живут так плохо. . .* — (*сказать, говорить*, прош. вр., жен. р., ед. ч.) *Наташа* (М. Горький. Мать, ч. I, гл. VI).

В соответствии с алгоритмом рассмотрим переводы. В испанском переводе (А. Эррайса и Х. Венто) находим имперфект:

— *Para entender por qué la gente vive tan mal. . . — d e c í a Natasha.*

Имперфект находим также во французском переводе С. Перского — *disait*, в румынском переводе (под ред. Т. Аркипа) — *spunea*, в румынском переводе Ч. Петреску и Н. Гане — *glăsuia*, в новогреческом переводе — ἔλεγε.

Согласно алгоритму, это означает, что здесь имеется означаемое, требующее в русском языке употребления н е с о в е р ш е н н о г о в и д а. Действительно, у М. Горького находим:

— *Чтобы понять, отчего люди живут так плохо. . . — говорила Наташа.*

Всегда, когда в романских языках употребляется имперфект, по-русски мы находим глагол несовершенного вида (но обратное неверно). Это значит, что семантическое пространство, соответствующее имперфекту романских языков, составляет часть семантического пространства русского несовершенного вида (прошедшего времени).

С т р о к и 2—9. Значительно сложнее правила определения вида глагола авторской речи при аористах в романских языках, греческом и урду (аористом будем для краткости называть как аорист в новогреческом языке, так и passé simple во французском, passato remoto в итальянском, pretérito indefinido в испанском, perfect simplu в румынском, простое прошедшее в урду).

С т р о к а 2. Если глагол авторской речи — *продолжать/продолжить*, то практически всегда он выступает в несовершенном виде:

— *И тем не менее, — п р о д о л ж а л бакалавр. . .* («Дон Кихот», перевод под ред. Б. А. Кржевского и А. А. Смирнова, т. II, гл. III).

В других переводах «Дон Кихота», где употреблены другие глаголы, на этом месте находим совершенный вид: *с к а з а л бакалавр* (перевод Любимова), *о т в е т и л бакалавр* (перевод М. Ватсон), *о т в е т и л Караско* (перевод Карелина). В испанском оригинале и французском переводе здесь употреблены аористы.

— *Одним из недостатков этой истории, — п р о д о л ж а л бакалавр, — считается то, что автор вставил в нее повесть под названием «Безрассудно-любопытный»* («Дон Кихот», перевод Любимова, т. II, гл. III). В других переводах наблюдаем такую же картину: если использован глагол *продолжить, -ать*, он употребляется в несовершенном виде (Кржевский: *п р о д о л ж а л бакалавр*, Карелин: *п р о д о л ж а л бакалавр*), если же употреблен другой глагол, находим совершенный вид (Ватсон: *с к а з а л бакалавр*). В текстах на романских языках здесь (как и в предыдущем примере) употреблен аорист, а не имперфект.

— *Я говорил, — п р о д о л ж а л Павел, — не о том добром и милостивом боге, в которого вы верите, а о том, которым попы грозят нам, как палкой* (М. Горький. Мать, ч. I, гл. XI). Ср. в испанском переводе аорист *continuó*.

Во всех исследованных нами литературных текстах (оригинальных и переводных) глагол *продолжить, -ать* при прямой речи выступает только в форме несовершенного вида. Вероятность несовершенного вида здесь практически равна 100%. Форма совершенного вида *продолжил* в принципе тоже возможна, но употребляется исключительно редко:

— *Время идет, и люди растут. — Анка с хитринкой поглядела на пастуха, п р о д о л ж и л а: — И ум у людей растет* (Г о р б у н о в. Ледолом, II, 7 — пример этого редкого употребления заимствован из «Словаря современного русского литературного языка», т. XI, слово «продолжать»).

Здесь употребление совершенного вида вызвано, видимо, синтаксическими условиями: при бессоюзном сочинении однородных сказуемых совершенный вид первого глагола-сказуемого, по-видимому, требует совершенного вида второго глагола.

Строки 3 и 7. Если глагол авторской речи — *ответить/отвечать*, то на выбор вида оказывает влияние прагматическая характеристика текста (стиль и время создания).

В произведениях XX в. на современную тематику совершенный вид в этих условиях употребляется с вероятностью, близкой к 100%. В частности, в исследованных произведениях М. Горького, Л. Андреева, К. Паустовского везде в этих случаях употреблен совершенный вид. У А. П. Чехова также совершенный вид *ответил* употребляется почти всегда, *отвечал* — редко. Вот один из редких случаев:

Потом он шагнул к костру, улыбнулся еще светлее и сказал:

— Хлеб да соль, братцы!

— Милости просим! — отвечал за всех Пантелей (А. П. Чехов. Степь, гл. VI).

В другом классе текстов — в современных (XX в.) переводах западноевропейской художественной прозы XV—XVIII вв. (Сервантес, Фильдинг, Дефо, Ш. де Костер) — употребление форм *отвечал* и *ответил* (при аористных временах в текстах на романских языках) колеблется в зависимости от индивидуального стиля переводчика, но вероятность употребления формы *отвечал* обычно выше, чем в произведениях современной тематики: у Н. Любимова («Дон Кихот») и А. А. Франковского (Г. Фильдинг. История Тома Джонса Найденыша и Д. Дефо. Молль Флендерс) частотность несовершенного вида составляет 70% и выше, у А. Горнфельда (Ш. де Костер, Легенда об Уленшпигеле) и у Н. Вольпин (Г. Фильдинг: «История жизни покойного Джонатана Уайльда Великого» и «История приключения Джозефа Эндруса») — только 10—20%. Есть и такие тексты переводов, где форма *отвечал* в этой функции совсем не встречается: таков текст «Дон Кихота» в переводе М. Ватсон и перевод «Дон Кихота» под редакцией Б. Кржевского и А. Смирнова. Следует, однако, отметить одно существенное обстоятельство: такие переводы воспринимаются как стилистически несовершенные или во всяком случае менее характерные для стиля переводов западной литературы XV—XVIII вв. и приближающиеся к «нейтральному» стилю.

В отношении форм, описанных в строках 3 и 7 алгоритма, можно поразному применять термины «совершенный вид» и «несовершенный вид»:

1. Учитывая, что форма *отвечал* используется в условиях, аналогичных условиям употребления совершенного вида других глаголов говорения (*сказал, возразил, спросил, отозвался* и др.), можно называть *отвечал* особой формой совершенного вида, омонимичной форме несовершенного вида *отвечал* и синонимичной форме совершенного вида *ответил*. Так, видимо, рассматривается форма *отвечал* в «Словаре современного русского языка»:

«Отвечать, аю, аешь, несов.; ... Отвечать, аю, аешь, в знач. сов. — *А, Федя! дома Хорь? — спросил его г-н Полутыкин. — Нет. Хорь в город уехал, — отвечал парень.* Тург. Хорь и Калиныч» (т. VIII, стр. 1275).

2. Можно считать форму *отвечал* в описанных условиях особым случаем употребления несовершенного вида. Последний способ обозначения принят в настоящей работе.

Оба способа обозначения в принципе равноправны и различаются лишь номенклатурно и, может быть, с точки зрения экономии описания. При первом способе обозначения строки 3 и 7 устраняются из нашего алгоритма, но зато приходится создавать особый сложный алгоритм для различения омонимии видов в форме *отвечал*.

Строка 5. Здесь рассматривается обычный случай: если в текстах на романских языках и на новогреческом употреблен аорист, то глагол при прямой речи, имеющий форму прошедшего времени, в русском языке — совершенного вида.

Исключения из этого правила рассмотрены в строках 6 и 8.

Примеры к строке 5 (глагол *ответила* иллюстрирует употребление, отмеченное в строке 7):

— *В субботу у меня будут гости из города.*

— *Из города? — повторила мать и вдруг всхлипнула.*

— *Ну, о чем, мамаша? — недовольно воскликнул Павел.*

Она, утирая лицо фартуком, ответила; вздыхая:

— *Не знаю, так уж...* (М. Горький. Мать, ч. I, гл. V).

Во всех тех языках, где проводится различие между аористом и имперфектом, здесь находим аорист. Например, во французском переводе (С. Перский):

— *Il viendra des gens chez moi, le samedi! ... de la ville.*

— *De la ville...? répéta la mère en hochant la tête.*

Soudain, elle se mit à sangloter.

— *Pourquoi pleurer, maman? s'écria Pavel mécontent. Pourquoi?*

Elle répondit d'une voix faible en essuyant ses larmes:

— *Je ne sais pas... comme cela...*

В итальянском переводе:

— *Sabato verranno de me certe persone...: dalla città.*

— *Dalla città? — chiese la madre..., e improvvisamente ruppe in singhiozzi.*

— *Che cosa vuol dire, mamma? — chiese lui con impazienza.*

Ella, asciugandosi gli occhi col grembiule, rispose a bassa voce:

— *Non lo so... Così!...*

В испанском переводе:

— *El sábado vendrá gente de la ciudad a verme.*

— *¿ De la ciudad? — repitió la madre, y de pronto empezó a sollozar.*

— *¿ Cómo es eso, madre? — exclamó Pável disgustado.*

Ella, enjugándose las lágrimas con el delantal, contestó suspirando:

— *No sé, porque sí...*

В румынском переводе (под ред. Т. Аркипа):

— *Sîmbătă au să vină la mine musafiri din oraş.*

— *Din oraş? — repetă mama, şi deodată o podidi plînsul.*

— *De ce, mamă? — strigă Pavel nemulţumit.*

Ea îşi şterse faţa cu şorţul şi răspunse oftînd:

— *Nu ştiu, aşa...*

В румынском переводе (Ч. Петреску, Н. Гане):

— *Sîmbătă îmi vin nişte oaspeţi din oraş.*

— *Din oraş? — repetă mama şi deodată o podidi plînsul.*

— *Dar ce este, mamă? — întrebă Pavel cu nemulţumire.*

Ştergîndu-şi faţa cu şorţul, ea răspunse suspinînd:

— *Nu ştiu, aşa...*

В новогреческом переводе:

—Θάρθουν ἄνθρωποι στό σπίτι, τό Σάββατο... ἀπ' τήν πόλη.

—'Απ' τήν πόλη... **ξανᾶπε** ἐκεῖνη, κουνώντας τό κεφάλι της. Καῖ ξαφρικά ἔβαλε τά κλάματα.

—Γιατί κλαῖς, μητέρα; τῆς **φώναξε** δυσαρεστημένος ὁ Πάβελ. Γιατί;

'Απάντησε μ'ἀδύνατη φωνή, σκουπίζοντας τά δακριά της:

—"Ετσι... Δέν ξέρω...

В переводе на хинди (Мунша Саксена):

«*Sanīcar ko šahar se kuch log mujhase milane āyeṅge*»

«*Šahar se?*» *māṅ ne usake šabda dohārāye aur sahsā vah rone lagī.*

«*Kyā bāt hai māṅ?*» *Pavel ne circirākar kahā.*

Māṅ ne apne dāman se āṅsu poṅch ḍāle.

«*Mālum nahīṅ*» *usne āh bharkar kahā.*

В переводе на урду:

«Sanīcar ko šahar se kuch log mujha se milane āyenge».

«Šahar se!» usa ki mān ne d o h a r ā y a aur phir dafatā vah siskiyān bharne lagī.

«Mān bāt kyā hai?» Pavel ne jhālkar p ū c h ā.

Us ne apne pīš band se ānsū ponche.

«Mujhe nahīn mālum» us ne thandā sāns bharte hue k a h ā.

Во всех этих переводах присутствуют аористы, которым в русском тексте, согласно алгоритму, соответствует совершенный вид глагола.

Строка 6. Рассмотрим такие высказывания:

Она воротилась поздно вечером, усталая, но довольная:

— Сашеньку видела! — г о в о р и л а она сыну. — Кланяется тебе. А этот Егор Иванович простой такой, — шутник! Смешно говорит. (М. Горький, Мать, ч. I, гл. XII);

Мать чувствовала, что от нее чего-то хотят, ждут, и торопливо с п р а ш и в а л а:

— Ну, так что же? Как же? (М. Горький. Мать, ч. I, гл. XIV);

Вместо того, чтобы обсыпать столб заслуженными ругательствами, Гараська обратился к нему с кроткими упреками, носившими несколько фамильярный оттенок.

— Стой, дурашка, куда ты? — б о р м о т а л он, откачиваясь от столба и снова всей грудью припадая к нему (Л. Андреев. Баргамот и Гараська);

— И — курить буду! Дай мне отцову трубку... — тяжело двигая непослушным языком, б о р м о т а л Павел (М. Горький. Мать, ч. 1, гл. III).

Если подставим в эти фразы *сказала* вместо *говорила*, *спросила* вместо *спрашивала*, *пробормотал* вместо *бормотал*, получим предложения, вполне допустимые в русском языке, но, по мнению носителей русского языка, отличные по значению от фраз М. Горького и Л. Андреева. Это различие, релевантное для русского языка, не отражается, однако, на переводах (по крайней мере, на тех переводах Горького, которые были исследованы). Например, глаголу *бормотал* в последнем примере в переводах соответствуют аористы: франц. *grogna*, исп. *barbotó*, ит. *borbottò*, рум. *mormăi*, греч. μουρμούρισε, урду *burburāyā*.

Возможно, что с привлечением переводов еще на какие-нибудь языки и — что вероятнее — при рассмотрении различных возможных переводов на те же языки удалось бы зафиксировать эти различия переводным методом. Можно попытаться также — хотя бы для части высказываний — искать зависимости между выбором вида и контекстом (хотя во всех исследованных пока переводах при наличии аористов во всех переводах несовершенный вид в русском тексте всегда заменим на совершенный с сохранением грамматической правильности и осмысленности фразы). Пока, однако, ничего этого не сделано.

Поэтому приходится прибегнуть к крайнему средству, принципиально нежелательному: задавать информанту вопросы, выходящие за пределы установления правильности фразы. Информантам задается вопрос: «Судя по тексту, автор привел все слова героя целиком или приведенные слова — только отрывок из речи героя?». Для части высказываний пригоден также вопрос: «Повторил ли герой приведенные слова несколько раз?» (например, для фраз типа: *Больно,—стонал он*; *Пожар! — кричал он*). Если большинство информантов считает, что приведенные слова — не отрывок, а вся речь героя (и герой произнес эти слова один раз, а не повторял их), тогда высказывание относится к классу, представленному строкой 5, и должно содержать совершенный вид. В случаях же, когда большинство информантов считает приведенную прямую речь отрывком, служащим для иллюстрации речи героя (в частности, герой мог повторять приведенные слова многократно), относим высказывание к классу, представленному стро-

кой 6 (в таких высказываниях по-русски употребляется несовершенный вид).

Строка 8. Если мнения информантов, получивших указанный выше вопрос, расходятся или значительная часть (более половины) информантов допускает обе возможности (например, отвечает «не знаю»), высказывание относим к классу 8 (случаи, когда возможно употребление обоих грамматических видов глагола, притом замена одного вида другим, по-видимому, не влияет вполне определенным образом на смысл высказывания).

К этому классу относятся (судя по предварительному, недостаточно широкому эксперименту) такие высказывания:

Узнав новость, он быстро вскочил, его лицо побледнело.

— Вот те и раз... — пробормотал он.

— Что надо делать-то? — дрожащей рукой отирая с лица пот, спрашивала Власова.

— Погодите, — вы не бойтесь! — ответил Федя, поглаживая здоровой рукой курчавые волосы (М. Горький. Мать, ч. I, гл. IX).

Заметим, что в переводах здесь повсюду употреблен не имперфект, а аорист (в соответствиях русскому глаголу *спрашивала*): франц. *demanda*, исп. *preguntó*, рум. *întrebă* (в обоих переводах), новогреч. ρώτησε, ит. *domandò*, урду *daryāft kiyā*.

Другие примеры.

Вот он вылез из воды и, держа ведро обеими руками, побежал к подводам.

— Уже полное! — кричал он, тяжело дыша. — Давайте другое! (А. П. Чехов. Степь, гл. V);

Тяжело крякнув..., Баргамот присел на корточки возле Гараськи.

— Ну... — смущенно гудел он. — Может, оно не разбилось?

— Да, не разбилось, ты и морду-то всю готов разбить. Ирод! (Л. Андреев. Баргамот и Гараська).

Следует заметить, что по мере увеличения количества переводов список высказываний, предварительно отнесенных к строке 8, сокращается в пользу строки 1 (так как среди новых переводов обнаруживались имперфекты, а по условиям алгоритма достаточно одного имперфекта в переводе, чтобы отнести высказывание к строке 1). Это еще более укрепляет надежду на возможность устранения необходимости обращаться к информанту с вопросами, выходящими за пределы установления отмеченности высказываний.

Судя по предварительным нашим наблюдениям, употребление несовершенного вида, описанное в строках 6 и 8, ограничено рамками текстов определенной прагматической характеристики: художественная проза XX в.[3] (оригинальная и переводы современных авторов), но преимущественно в жанрах романа и повести (гораздо реже рассказа и очерка[4]). Такое употребление, по-видимому, совершенно отсутствует в переводах западных писателей XV—XVIII вв. и античности (мы не обнаружили этого употребления в переводах произведений Сервантеса, Дидро, Фильдинга, Ш. де Костера, Апулея).

Такое ограничение подобного употребления несовершенного вида допускает следующее толкование. Это употребление связано с определенным литературным приемом: речь героя изображается не путем полного цитирования, а путем намеренного приведения отрывков, иллюстрирующих речь героя. Такой прием появляется в литературе,

[3] Проза XIX в. не исследовалась.

[4] Это употребление обнаружено в романе М. Горького «Мать», в повести А. П. Чехова «Степь», но не обнаружено в рассказах М. Горького, А. П. Чехова, А. И. Куприна, К. Г. Паустовского, В. М. Гаршина, а также в сборнике переводов современных авторов «Зарубежная новелла» (М., 1962).

видимо, довольно поздно и поэтому обычно не встречается при переводах авторов, относящихся к той эпохе, когда этот прием не был распространен.

Строка 9. Глагол (*по*)*думать* при прямой речи (в случаях, когда в переводах употреблен аорист) весьма часто выступает в форме несовершенного вида (не реже, а может быть, и чаще, чем в совершенном виде):

— *Вот кого потревожили!* — *думала она* (М. Горький. Мать, ч. I, гл. X). В переводах на языки, различающие аорист и имперфект, здесь находим аорист: франц. *pensa*, исп. *pensó*, ит. *pensò*, рум. *gîndi* (в обоих переводах), новогреч. σκέφτηκε, урду *socā*.

Строка 10. Плюсквамперфекту западных языков в русском языке при прямой речи соответствует как несовершенный вид, так и совершенный:

— *Я зна-а-ю, как это получится,* — *говорила ему сестра, когда он готовился к отъезду в деревенскую глушь* (Саки /Гектор Монро/. Открытое окно, перевод с англ., в сб. «Зарубежная новелла»).

В этом случае по-русски можно было бы заменить *говорила* на *сказала*, а грамматическая правильность фразы (и ее осмысленность) сохранились бы.

Пока не выяснено, какой вид глагола здесь употребляется чаще и какие закономерности определяют выбор вида в тех случаях, которым в западных языках соответствует плюсквамперфект.

Рассмотренная область употребления — глаголы при прямой речи — составляют лишь очень малую часть всего круга употреблений видов в русском языке, но алгоритм, описывающий это употребление, может, вероятно, служить образцом для построения алгоритма, описывающего употребление вида в целом.

Список использованных текстов

Леонид Андреев. Повести и рассказы. М., 1957.
В. М. Гаршин. Сочинения. М.—Л., 1951.
М. Горький. Мать (М. Горький. Собрание сочинений. Изд. 3, т. IV (1941).
М. Горкī. Māṅ. Māsko, 1950 (хинди, переводчик — Мунīшā Саксена).
М. Горки. Māṅ. Māsko, [б. г.] (урду).
M. Gorki. La madre. Mosca, 1943.
M. Gorki. La madre. Traducción del ruso A. Herraiz y J. Vento. Moscú, 1951.
M. Gorchi. Mama. Ed. 4-a revăzută de T. Archip. Bucureşti, 1952.
M. Gorki. Mama. Trad. de Cesar Petrescu şi Nicolae D. Gane. Bucureşti, 1955.
M. Gorki. La mère. Trad. S. Persky. Paris, 1907.
M. Gorky. Mother. Translated from the Russian by Margaret Wettlin. Moscow, 1954.
Μ. Γκόρκι 'Η μητέρα. 'Εκδ. «Νέα 'Ελλάδα», 1951.
М. Горький. Собрание сочинений. Изд. 3, т. IV. М., 1941 (рассказы).
Даниель Дефо. Молль Флендерс. Пер. А. Франковского. М., 1955.
«Зарубежная новелла». М., 1962.
Шарль де Костер. Легенда об Уленшпигеле и Ламме Гудзаке, их приключениях отважных, забавных и достославных во Фландрии и иных странах. Пер. А. Горнфельда. М., 1955.
А. И. Куприн. Избранные сочинения. М., 1947.
К. Г. Паустовский. Избранные произведения в двух томах. М., 1956.
М. Сервантес Сааведра. Дон-Кихот Ламанчский. Пер. с испанского В. Карелина, исправленный и дополненный В. Зотовым. М., 1910.
М. де-Сервантес Сааведра. Остроумно-изобретательный идальго Дон-Кихот Ламанчский. Пер. М. В. Ватсон, ч. I—II. Пг., 1917.
М. де Сервантес Сааведра. Хитроумный идальго Дон Кихот Ламанчский. Пер. под ред. Б. А. Кржевского и А. А. Смирнова. Л., 1949.
М. де Сервантес Сааведра. Хитроумный идальго Дон Кихот Ламанчский Пер. Н. Любимова, ч. 1—2. М., 1959.
M. de Cervantes Saavedra. El ingenioso hidalgo Don Quijote de la Mancha (по изданию: Buenos Aires, 1954).
M. de Cervantes Saavedra. The achievements of the ingenious gentleman Don Quixote de La Mancha. A translation based on that of Peter Anthony Motteux, witx the memoir and notes of John Gibson Lockhart. Vol. I—II. London, 1891.

M. de Cervantes Saavedra. Leben und Taten des scharfsinnigen Edlen Don Quijote von La Mancha. Übersetzt von Roland Schacht. Bd. 1—2. Potsdam, 1951.
Histoire de l'admirable Don Quixotte de Ia Manche. T. III. Amsterdam, 1700.
Г. Фильдинг. Избранные произведения в двух томах. М., 1954.
А. П. Чехов. Рассказы и повести, М., 1944.

VERBAL ASPECT IN RUSSIAN AND THE STOCHASTIC DEPENDENCE OF THE *SIGNIFIANT* UPON THE *SIGNIFIÉ*

Summary

If one knows the context of a verbal form and its *signifié*, its aspect may be found out by means of a certain algorithm. In the present paper the author describes a part of such an algorithm.

The algorithm is presented in the form of a table containing seven columns. In the first six columns we record the information necessary for obtaining the answer; the seventh column contains the answer thus obtained. Columns 1—4 contain a description of the context of a verbal form. Columns 5 and 6 pertain to the *signifié* and contain (5) translation of the form in question and (6) additional information. Column 7 furnishes the answer we seek for by defining the aspect of a verbal form as perfective or imperfective or pointing out the probability of this or that aspect if both aspects are possible in the given context and with the given *signifié*.

The verbal aspect in Russian illustrates the fact that there is a stochastic dependence of the *signifiant* upon the *signifié*.

БИБЛИОТЕКА
ИН-ТА РУССКОГО ЯЗЫКА
ИН-ТА ЯЗЫКОЗНАНИЯ АН СССР

СОДЕРЖАНИЕ

I. ОБЩИЕ ВОПРОСЫ

С. К. Шаумян. О логическом базисе лингвистической теории 3
S. K. Šaumjan. On the logical basis of linguistic theory (Summary) 8
А. А. Зиновьев. Об основах абстрактной теории знаков 9
A. A. Zinovjev. On the grounds of abstract theory of signs (Summary) . . . 22
Ю. К. Лекомцев. Элементы теории языковой сочетаемости 23
Y. K. Lekomtsev. Elements of a theory of distribution of language units (Summary) 45
С. Маркус. Логический аспект лингвистических оппозиций . . . . .
S. Marcus. The logical aspect of linguistic oppositions (Summary) . . . . .
Б. В. Сухотин. Алгоритмы лингвистической дешифровки 75
B. V. Sukhotin. Algorithms for linguistic deciphering (Summary) 101
Ю. Д. Апресян. Современные методы изучения значений и некоторые проблемы структурной лингвистики 102
Y. D. Apresyan. Modern methods in the study of meaning and some problems of structural linguistics (Summary) 149

II. ФОНОЛОГИЯ

А. Росетти. Значение спектрального анализа для фонологии . . . . .
A. Rosetti. The value of the spectrograph analysis in phonology (Summary) . 152
Л. И. Богораз. О нейтрализации и архифонеме в связи с согласными архифонемами русского языка 153
L. I. Bogoraz. On the concepts of «neutralization» and «archiphoneme» in connection with the consonant archiphonemes in Russian (Summary) 163
В. В. Шеворошкин. О структуре звуковых цепей 164
V. V. Ševoroškin. On the structure of sound-chains (Summary) 181
Е. Л. Гинзбург. Некоторые случаи неодносоставности просодемоидов . . . 182
E. L. Ginsburg. On some cases of non-unitariness of prosodemoids (Summary)
Э. Василиу. Структура консонантных групп в румынском языке 210
E. Vasiliu. The structure of consonant groops in Rumanian (Summary) 219

www.ruslang.ru

III. ГРАММАТИКА И ЛЕКСИКОЛОГИЯ

И. И. Ревзин. Опыт символического изложения немецкого синтаксиса применительно к проблемам машинного перевода 220
I. I. Revzin. A simbolization of German syntax and its application to problems of mechanical translation (Summary) 232
П. А. Соболева. Опыт исчисления трансформаций на основе теории С. К. Шаумяна о порождении классов слов в процессе порождения грамматики 233
P. A. Soboleva. Experiment of transformation calculus on the basis of Šaumjan's theory of word class generation in the process of grammar generation (Summary) 265
А. Б. Долгопольский. Категория вида в русском языке и вероятностный характер связи означаемого с означающим 266
A. B. Dolgopolsky. Verbal aspect in Russian and the stochastic dependence of the *signifiant* upon the *signifié* (Summary) 281

Проблемы структурной лингвистики

Утверждено к печати
Институтом русского языка Академии наук СССР

РИСО № 80—125В. Сдано в набор 10/VI 1963 г. Подписано в печать 26/IX 1963 г. Формат 70×108¹/₁₆. Печ. л. 17,75=24.31 усл. печ. л. Уч.-изд. л. 22,8. Тираж 3000. Изд. № 1937. Тип. зак. 248. *Цена 1 руб. 47 коп.*

Издательство Академии наук СССР
Москва, К-62, Подсосенский пер., 21

1-я типография Издательства АН СССР.
Ленинград, В-34, 9-я линия, д. 12

ОПЕЧАТКИ

Страница	Строка	Напечатано	Должно быть
53	2 св.	падеж, = степень	падеж, степень
55	1 »	синий, черный	зеленый, черный
55	3 »	$A_2 - B_2$, $B - A$	$A_2 - B_2$, $B_1 - A_1$
82	2 сн.	*Б*	*АБ*
112	2 св.	истинности языка-объекта	истинности высказываний языка-объекта
123	7—8 св.	«истинное (допускать, что *x* истинно) определяет	«истинное» определяет
140	19 св.	по активности	и активности
142	26—27 св.	множества	множеств
143	17 св.	языка	языки

Проблемы структурной лингвистики. 1963

1 р. 47 к.