4
П1-71/30
18/II

ИНСТИТУТ РУССКОГО ЯЗЫКА АН СССР

Проблемная группа по экспериментальной и прикладной лингвистике

ПРЕДВАРИТЕЛЬНЫЕ ПУБЛИКАЦИИ

Выпуск 30

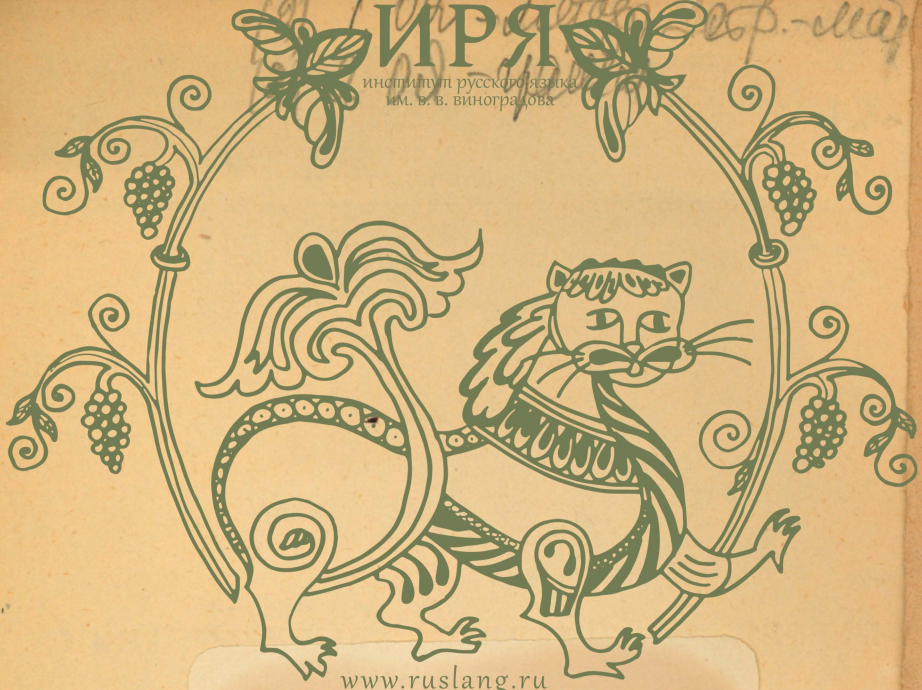
И. А. МЕЛЬЧУК

I. УРОВНИ ПРЕДСТАВЛЕНИЯ ВЫСКАЗЫВАНИЙ И ОБЩЕЕ СТРОЕНИЕ МОДЕЛИ «СМЫСЛ ⟷ ТЕКСТ»

II. СЛОВООБРАЗОВАНИЕ И КОНВЕРСИЯ

Москва — 1972

ИРЯ
институт русского языка
им. в. в. виноградова



www.ruslang.ru

66823

ИНСТИТУТ РУССКОГО ЯЗЫКА АН СССР

Проблемная группа по экспериментальной
и прикладной лингвистике

ПРЕДВАРИТЕЛЬНЫЕ ПУБЛИКАЦИИ

Выпуск 30

И.А.МЕЛЬЧУК

I. УРОВНИ ПРЕДСТАВЛЕНИЯ ВЫСКАЗЫВАНИЙ
И ОБЩЕЕ СТРОЕНИЕ МОДЕЛИ "СМЫСЛ ⟷ ТЕКСТ".

II. СЛОВООБРАЗОВАНИЕ И КОНВЕРСИЯ.

Москва-1972

4
П-7/30
БИБЛИОТЕКА
ИН-ТА РУССКОГО ЯЗЫКА
ИН-ТА ЯЗЫКОЗНАНИЯ АН СССР
66823

В связи с исследованиями, которые проводятся членами Проблемной группы экспериментальной и прикладной лингвистики (ПГЭПЛ) при ИРЯ АН СССР, в ряде случаев возникает необходимость достаточно отчетливо представлять себе ту общую концепцию естественного языка, к которой эти исследования в той или иной степени примыкают, - так называемую модель "Смысл ⟷ Текст". Особенно остро такая необходимость проявилась при разработке формальной модели русской морфологии (см. Еськова и др. 1971), в частности - во время I-ой (1970 г.) и II-ой (1972 г.) Зимних школ по экспериментальной и прикладной лингвистике. Поэтому осенью 1971 года в ИРЯ АН СССР было начато чтение лекций по модели "Смысл ⟷ Текст" - с целью хотя бы поверхностно ознакомить с ней научную общественность, в том числе - членов ПГЭПЛ.

Однако эти лекции не могут, разумеется, заменить о п и с а н и е модели, каким бы кратким и приблизительным оно ни было. Именно поэтому автор решил издать в серии предварительных публикаций ПГЭПЛ'а сжатое и схематичное описание модели "Смысл ⟷ Текст", которое могло бы служить временным справочным пособием и введением в соответствующую проблематику - д л я д о с т а т о ч н о п о д г о т о в л е н н и х ч и т а т е л е й.

Последние слова выделены не случайно.

С одной стороны, изложение в настоящей брошюре начинается ex abrupto : ничего не говорится о содержательных лингвистических соображениях, на которых основывается концепция моделей типа "Смысл ⟷ Текст", не разъясняется (применительно к языку) понятие функционального, или кибернетического, моделирования, и т.п.

С другой стороны, оно не доводится до естественного завершения: вводимые понятия лишь слегка намечены и должны, разумеется, рассматриваться вновь с соответствующей подробностью при систематическом описании всех компонентов модели; опущены многие важные разъяснения; оставлен ряд логических лакун.

Таким образом, многое в данной брошюре может быть не до конца понятным, чего, однако, не следует пугаться: предлагаемый ма-

териал призван играть роль как бы очень обобщенного конспекта или даже оглавления к полному описанию модели – и только. Автор будет считать свою цель достигнутой, если читатель брошюры получит из нее самое общее (пусть не всегда полностью отчетливое) представление о строении модели "Смысл ⇔ Текст" в целом[1].

Автор многим обязан ряду своих коллег, в особенности – А.К.Жолковскому, совместно с которым формулировались и разрабатывались идеи, нашедшие отражение в этой брошюре. Рукопись была прочитана М.К.Валиевым, А.Я.Диковским, А.К.Жолковским, Л.Н.Иорданской, Н.Г.Михайловой, Л.С.Модиной, Н.В.Перцовым, Б.Ю.Розенцвейгом и С.Я.Фитиаловым, сделавшими многочисленные полезные замечания, которые автор постарался учесть в меру своих возможностей.

Прежде всего, сформулируем вкратце те допущения, на которые опирается последующее изложение.

1) Естественный язык рассматривается как особого рода преобразователь "Смысл ⇔ Текст", а именно как отображение (бесконечного) множества смыслов на (бесконечное) множество текстов и обратно. Знать тот или иной конкретный язык – это значит уметь поставить в соответствие любому смыслу текст или равнозначные тексты на этом языке, а любому тексту – его смысл (или смыслы – если текст неоднозначен), и притом так, как это делают компетентные носители данного языка.

2) Язык-преобразователь можно представлять себе как совокупность весьма сложных правил, закодированных в мозгу носителей языка тем или иным способом; конкретная природа (физиологическая, химическая, ...) этого способа нас как лингвистов не интересует.

3) Понимаемый так язык недоступен для прямого, непосредственного наблюдения – во всяком случае, если пользоваться только лин-

[1] Для лучшей ориентации заинтересованного читателя в конце брошюры указаны некоторые публикации, так или иначе относящиеся к модели "Смысл ⇔ Текст".

вистическими методами. Тем самым, "чистому" лингвисту (не занимающемуся психологией или психофизиологией речи и т.п.) приходится создавать функциональные, или кибернетические, модели языка – с тем, чтобы, сравнивая результаты их функционирования с наблюдаемыми фактами речи, совершенствовать эти модели, все больше и больше приближая их к "настоящему" языку. Отсюда и возникает стремление описывать естественный язык, т.е. преобразователь "смыслы ⟷ тексты", с помощью функциональной модели типа "Смысл ⟷ Текст".

4) Общеизвестно, что в естественных языках широко распространены, с одной стороны, омонимия/полисемия, а с другой – синонимия, так что соответствие "смыслы ⟷ тексты" является многомногозначным, а в силу этого – чрезвычайно сложным. Построение правил прямого, непосредственного перехода от смыслов к текстам и обратно оказывается слишком трудным или даже вообще невозможным. Поэтому приходится вводить ряд промежуточных уровней между смыслом и текстом, так что интересующий нас переход распадается на несколько последовательных этапов (каждый из которых оказывается много проще и обозримее процесса в целом). Разумеется, промежуточные уровни следует вводить не как угодно, а сообразуясь с языковыми фактами – именно это позволяет добиваться упрощения преобразований.

5) В соответствии со сказанным в пункте 4 модель "Смысл ⟷ ⟷ Текст" состоит из нескольких последовательно работающих компонентов – таких, что n-ый компонент модели отвечает за переход между уровнями n и $n+1$, т.е. как бы переводит представление n-го уровня в представление $(n+1)$-го уровня – и наоборот.

Настоящее изложение разбито на два раздела:

– Раздел I содержит общую характеристику используемых уровней представления высказываний[1], в нем же – после описания уров-

[1] Под высказыванием понимается любой достаточно автономный речевой акт; таким образом, в общем случае высказывание по объему превосходит фразу.

ней и приведения примеров – разъясняется само понятие уровня представления.

– Раздел II посвящен общей характеристике основных компонентов модели "Смысл ⟷ Текст".

Поскольку, как уже было сказано, мы стремимся дать здесь понятие лишь о всей картине в целом, мы вынуждены допустить известный схематизм и приблизительность в формулировках, а также отказаться от ряда существенных пояснений.

РАЗДЕЛ I.

Мы будем различать пять уровней представления высказываний (считая от смысла – наиболее глубинный уровень – к тексту – наиболее поверхностный уровень):

I. Семантический (смысл).

II. Синтаксический, включающий два подуровня:

A. Глубинно-синтаксический;

B. Поверхностно-синтаксический.

III. Морфологический, также включающий два подуровня:

A. Глубинно-морфологический;

B. Поверхностно-морфологический.

IV. Фонологический.

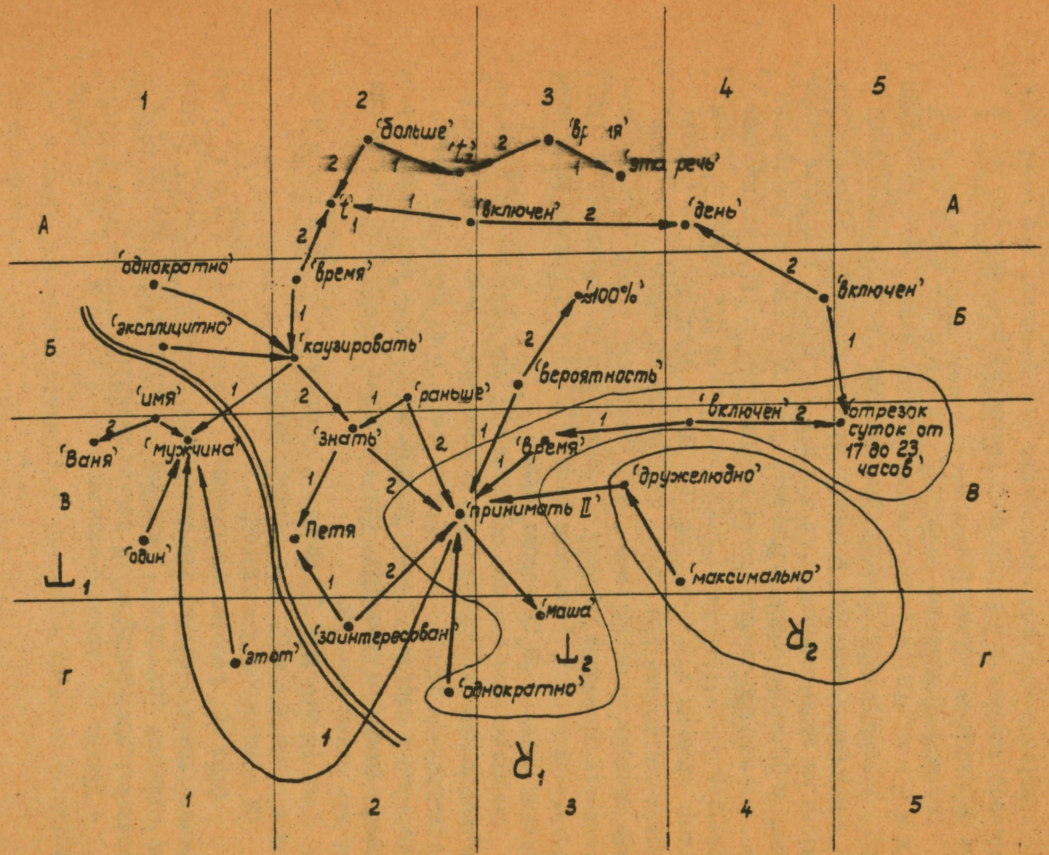
V. Фонетический/графический (текст).

Рассмотрим каждый из этих уровней.

I. На семантическом уровне целому высказыванию ставится в соответствие его семантическое представление (СемП); см. (1) – СемП русского высказывания Ваня твердо обещал Пете, что вечером он примет Машу самым тёплым образом[1]. (Разумеется, это же СемП отвечает и всем высказываниям, которые синонимичны данному.)

Семантическое представление есть запись высказывания на специальном (= построенном именно для этой цели) семантическом языке; оно состоит из двух следующих компонентов.

[1] Это СемП преследует, прежде всего, иллюстративные цели – и ни в малейшей степени не претендует на точность и окончательность.


1
2
3
4
5
А
Б
В
Г
'больше'
'вот я'
'эта речь'
'включен'
'день'
'однократно'
'время'
'включен'
'эксплицитно'
'каузировать'
'100%'
'вероятность'
'имя'
'раньше'
'включен'
'отрезок суток от 17 до 23 часов'
'Ваня'
'мужчина'
'знать'
'время'
'дружелюдно'
'принимать II'
'один'
Петя
'максимально'
'маша'
'этот'
'заинтересован'
'однократно'
(1)
-7-

4. С е м а н т и ч е с к и й г р а ф (СГ) – размеченный связный ориентированный граф, задающий смысл высказывания безотносительно к членению этого последнего на фразы и слова; такие конкретно-языковые особенности высказывания, как выбор определенной лексики и определенных синтаксических конструкций, в его СГ никак не отражается.

Вершины СГ помечены символами с е м а н т и ч е с к и х е д и н и ц (СЕ) – "слов" семантического языка, которые должны быть, вообще говоря, "смысловыми атомами", т.е. элементарными (= далее неразложимыми) в рамках данного описания; элементарные смыслы мы будем называть с е м а м и. Однако в порядке сокращения в СГ допускаются "промежуточные" СЕ – символы комплексов сем, т.е. целых семантических графов. Так, в (4), см. квадрат Б2, СЕ X ←1– 'раньше' –2→ Y – это сокращение для 'время t_2 Y-а больше времени t_1 X-а', ср. Б2, А2-3; аналогично СЕ 'знать' (В2) – сокращение для 'иметь информацию о', 'Маша' (Г3) – для 'эта женщина по имени Маша', а СЕ 'включен' – для 'быть подмножеством'; очевидным образом неэлементарны СЕ 'принимать ТР' (В2), 'дружелюбно' (В3), 'заинтересован' (Г2) и т.п. Полное разложение на семы сделало бы СГ необозримым.

Семантические единицы обозначаются для наглядности русскими словами в лапках; однако следует постоянно иметь в виду, что в СГ фигурируют не реальные русские слова с их многозначностью, конкретными синтаксическими и морфологическими свойствами и т.п., а некие условные этикетки, смысл которых должен быть всякий раз уточнен явным образом.

Дуги (= стрелки) СГ помечены символами с е м а н т и ч е с к и х с в я з е й, т.е. связей между СЕ.

СЕ делятся на два класса:

– Одни СЕ, условно называемые к о н с т а н т а м и, – это имена классов объектов или конкретных объектов ('поезд', 'скрепка', 'луна', 'человек', 'девушка', 'Светлана'); в эти СЕ дуги семантических связей могут только входить.

– Другие СЕ, условно называемые п р е д и к а т а м и, – это:

имена предикатов (‘плакать’, ‘широкий’, ‘выше’, ‘преобразовать’, ‘над’, ‘цель’, ‘часть’, ‘каузировать’, ‘максимально’, ...),

имена кванторов (‘все’, ‘существует’, ‘существует ровно n’, ...)

и имена логических связок (‘и’, ‘или’, ‘если’, ...).

Дуги семантических связей могут как входить в эти СЕ, так и выходить из них. Дуга-стрелка направляется от предиката к его аргументу.

Семы-предикаты – не более чем двухместны; промежуточные СЕ-предикаты могут иметь до пяти мест. Разные дуги, выходящие из одной вершины, всегда нумеруются разными номерами: $X \xrightarrow{1} \text{‘каузировать’} \xrightarrow{2} Y$, т.е. X – первый аргумент предиката ‘каузировать’ (т.е. X – причина: то, что ведет к / вызывает Y), а Y – его второй аргумент (Y – следствие или результат: то, что вызывается действиями или просто наличием X-а).

Мы будем считать СГ правильным, если 1) из каждого n-местного предиката в нем выходит не более n разных дуг, помеченных каждая номером места предиката, и 2) ни из одной константы не выходит никакая дуга.

2. Сведения о коммуникативной организации (КО) высказывания:

тема ($\perp$) – рема ($\exists$) [тема: то, о чем говорится в этом высказывании; рема – что именно сообщается о теме; в литературе приняты также названия: "актуальное членение", или "функциональная перспектива предложения"];

старое – новое [старое: то, о чем уже говорилось в предыдущих высказываниях; новое – то, что сообщается вновь];

эмфаза и т.п.

Соответствующие указания характеризуют определенные подграфы целого СГ; они относятся к СГ примерно так же, как просодические явления (ударение, интонация, ...) – к цепочке фонем, образующих фразу.

NB: в наших упрощенных примерах мы указываем сведения только о теме – реме.

Членение на тему и рему может иметь иерархические ярусы; так, в (1) внутри ремы первого яруса Y_1 имеется свое членение на I_2-Y_2.

Краткий комментарий к СемП (1).

1) 'Эксплицитно каузировать знать' = сообщать, давать знать; элемент 'эксплицитно' отличает это значение от дать понять = 'имплицитно каузировать знать'; ср. 'обещать' на стр. 25.

2) 'Время этой речи t_2 больше времени каузации t_1' (А2-3, Б2) = = 'каузация предшествует моменту речи', т.е. 'каузация' - в прошедшем времени; 'время' - название соответствующей шкалы, или координаты (как 'длина', 'скорость', 'вес', ...).

3) 'Однократно' (Б1 и Г2) ≈ совершенный вид соответствующих глаголов.

4) 'Знание о приеме раньше приема' (В2) ≈ принимать - в будущем времени.

5) 'Заинтересован' (Г2) ≈ его существенно касается (а не просто 'желает').

6) 'Отрезок суток от 17 до 23 часов' (Б5) = вечер [ср. в 4 часа дня - в 5, 6, ..., 10, 11 часов вечера - в 12 часов ночи].

7) 'Время Х-а включено в вечер' (В2-5) = Х - вечером.

8) ... говорил / сказал, что вечером примет - имеется в виду тот же день, когда говорит; поэтому 'время каузации знания' (В2) 'включено' (А2) в тот же день (А4), что и вечер (Б5), о котором идет речь.

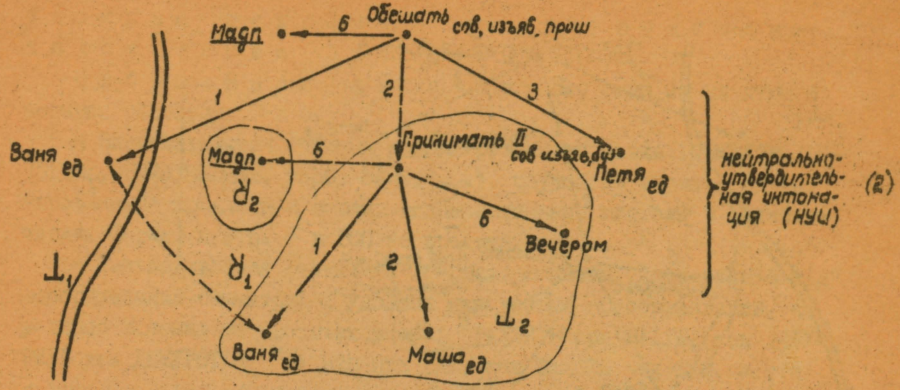
o

II. С и н т а к с и ч е с к и й у р о в е н ь.

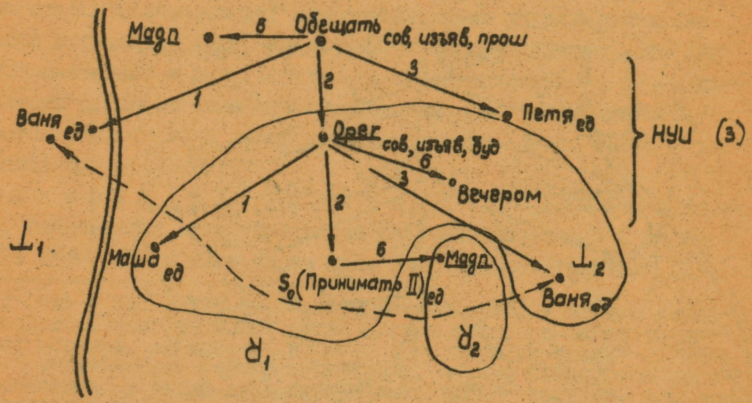
IIА. На уровне г л у б и н н о г о с и н т а к с и с а и на всех более поверхностных уровнях высказывание представляется пофразно; каждой отдельной фразе ставится в соответствие ее г л у б и н н о - с и н т а к с и ч е с к о е п р е д с т а в л е н и е (ГСП); см. (2)-(5) - ГСП четырех синонимичных фраз, отвечающих СемП (1).

ГСП фразы состоит из пяти следующих компонентов.

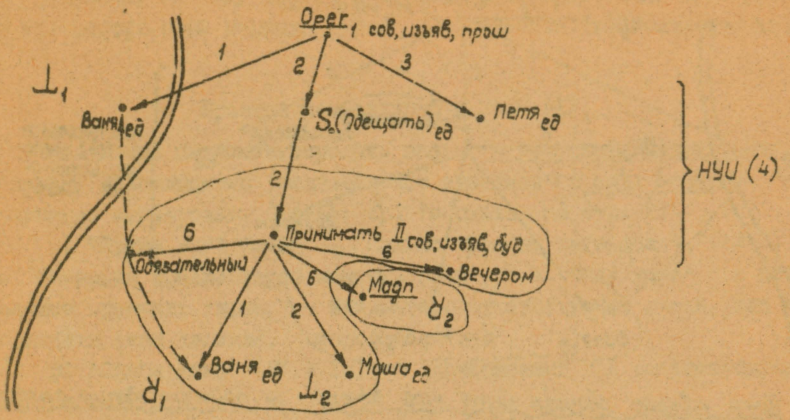
1. Г л у б и н н о - с и н т а к с и ч е с к а я с т р у к т у р а фразы (ГСС) - дерево зависимостей, узлы которого суть о б о-



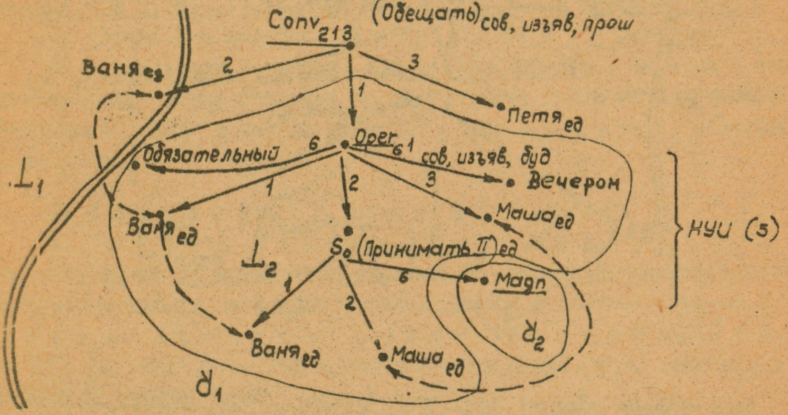
Ваня твердо обещал Пете, что вечером он примет Машу самым теплым <сердечным> образом.



Ваня твердо обещал Пете, что вечером Маша найдет у него самый теплый <радушный, сердечный> прием.



Ваня дал Пете обещание вечером обязательно принять Машу самым теплым образом.



Ване было обещано Пете, что вечером он обязательно окажет Маше самый теплый <радушный, сердечный> прием.

бщенные лексемы, а ветви - глубинно-синтаксические отношения (ГСО)[1].

Обобщенная лексема - это один из следующих четырех объектов:

- Знаменательная лексема рассматриваемого языка (ОБЕЩАТЬ, ВЕЧЕРОМ, ОБЯЗАТЕЛЬНЫЙ, ОКНО, ВАНЯ, ...). Служебные слова - сильноуправляемые предлоги и союзы, вспомогательные глаголы в аналитических формах и т.п. - в ГСС не представляются.

- Так называемая фиктивная лексема: лексема, "подсказываемая" соображениями симметрии и удобная для разных преобразований, но реально в языке не представленная: *ВАХТИТЬ при ВАХТА, как ДЕЖУРИТЬ при ДЕЖУРСТВО, *ГАМИТЬ при ГАМ, как ШУМЕТЬ при ШУМ, и т.п.

- Фразема определенного типа (≈идиома) вроде СЪЕСТЬ СОБАКУ, ИМЕТЬ ЗУБ, СИНИЙ ЧУЛОК, ВЫНОСИТЬ СОР ИЗ ИЗБЫ, ДАТЬ ПО ШАПКЕ,...

- Символ лексической функции ($Copy_{213}$, $Magn$, $Oper_1$, $Oper_2$, S_0, ...) , см. ниже.

При символе обобщенной лексемы необходимы индексы тех морфологических характеристик, которые несут смысл и не обусловлены синтаксически ("смысловыми" морфологическими характеристиками являются в русском языке, например, число существительного или время, наклонение и вид глагола - в противоположность таким синтаксически обусловленным характеристикам, как падеж существительного, род, число и падеж прилагательного, лицо, число и род глагола); см. в (2)-(5).

Лексическая функция (ЛФ) f - это зависимость, связывающая некоторое слово или словосочетание C (аргумент ЛФ) с множеством слов или словосочетаний $f(C)$

---

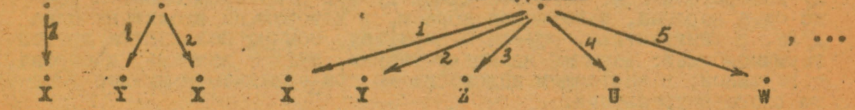
[1] Дерево - это связный ориентированный граф, в каждую вершину которого входит не более одной дуги и в котором имеется ровно одна вершина, называемая корнем, в которую не входит ни одна дуга. В лингвистике, говоря о деревьях, обычно пользуются другой терминологией: вершины называют узлами, дуги - ветвями, а корень - вершиной; в настоящем изложении мы придерживаемся именно такого словоупотребления.

(значение ЛФ) таким образом, что для любых C^1 и C^2, если только $f(C^1)$ и $f(C^2)$ существуют, то по смыслу и синтаксической роли в тексте $f(C^1)$ относится к C^1 точно так же, как $f(C^2)$ к C^2, причем для разных C в большинстве случаев f (C) тоже различны; это значит, что f(C) фразеологически связано с C. Выделено около 50 стандартных элементарных ЛФ – таких, у которых число возможных значений аргумента и число возможных значений функции достаточно велико; используются также сложные ЛФ, составляемые из стандартных элементарных. Примеры ЛФ: 1) синоним: Syn (приглашать) = звать, Syn (худой) = тощий; 2) конверсив: $Conv_{21}$ (спереди) = сзади, $Conv_{21}$ (следовать) = предшествовать, $Conv_{21}$ (обещать) = быть обещанным; 3) S_0 (обещать) = обещание, S_0 (посещать) = посещение, S_0 (препираться) = препирательства; 4) Magn (обещание / обещать) = твердое // твердо, Magn (беречь) = как зеницу ока, Magn (белый) = снежно=, ослепительно [≈ очень]; 5) $Oper_1$ (приказ) = давать, отдавать, $Oper_1$ (прием II) = оказывать, $Oper_1$ (мнение) = иметь, придерживаться [≈ 'быть субъектом']; 6) $Oper_2$ (приказ) = получать, $Oper_2$ (прием) = находить, встречать [≈ 'быть объектом']; 7) $Real_1$ (обещание) = сдерживать [≈ 'выполнять, будучи субъектом']; 8) $Real_2$ (приказ) = выполнять [≈ 'выполнять, будучи объектом'] ; и т.п.

ЛФ играют важную роль с точки зрения описания ограниченной лексической сочетаемости и в синонимическом перифразировании.

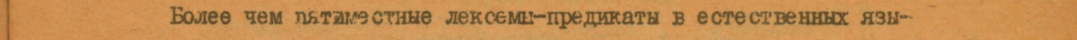
Глубинно-синтаксическое отношение – это одно из следующих восьми отношений:

ГСО 1–5 связывают лексему-предикат с ее первым, ..., пятым аргументом соответственно:

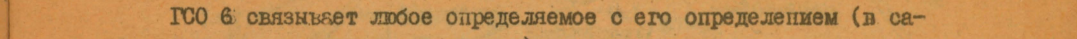


Более чем пятиместные лексемы-предикаты в естественных языках пока не известны.

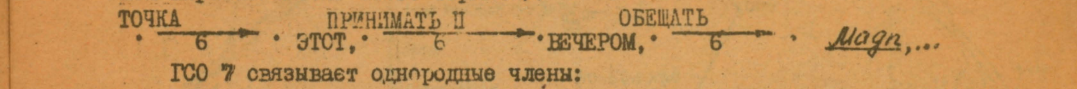
ГСО 6 связывает любое определяемое с его определением (в самом широком смысле этого термина):



ГСО 7 связывает однородные члены:



ГСО 8 присоединяет обращение, междометия и т.п. к вершине фразы:



сов, взяв, буд. (= Ваня, Маша придет).

№ : Линейный порядок узлов ГСО не задается! Тем самым структурные, иерархические отношения ("что с чем синтаксически связано и как именно") последовательно отделяются от отношений совсем иной природы – от отношений линейного расположения единиц речи ("что справа/слева от чего"). Эти последние отношения фиксируются на глубинно-морфологическом уровне и всех более поверхностных уровнях (см. стр. 18).

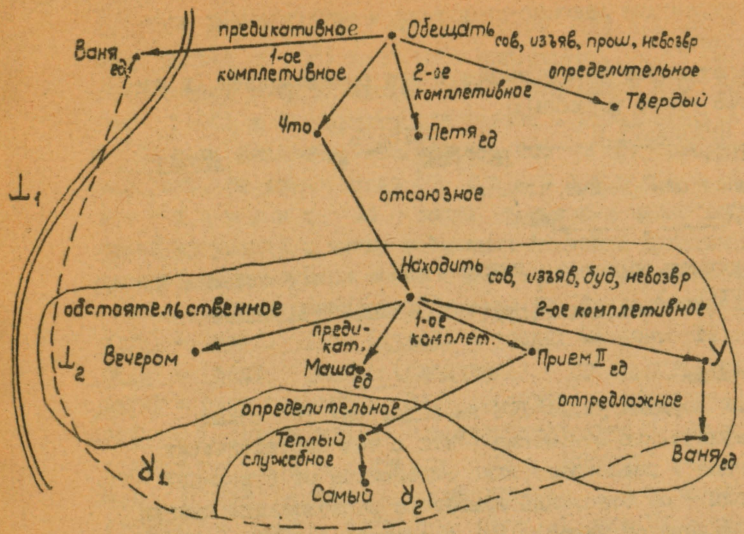
ГСО – центральный компонент глубинно-синтаксического представления; его прочие компоненты таковы:

2. Сведения о коммуникативной организации фразы – в частности, указание темы и ремы, см. ⊥ и ∂ в (2)-(5).

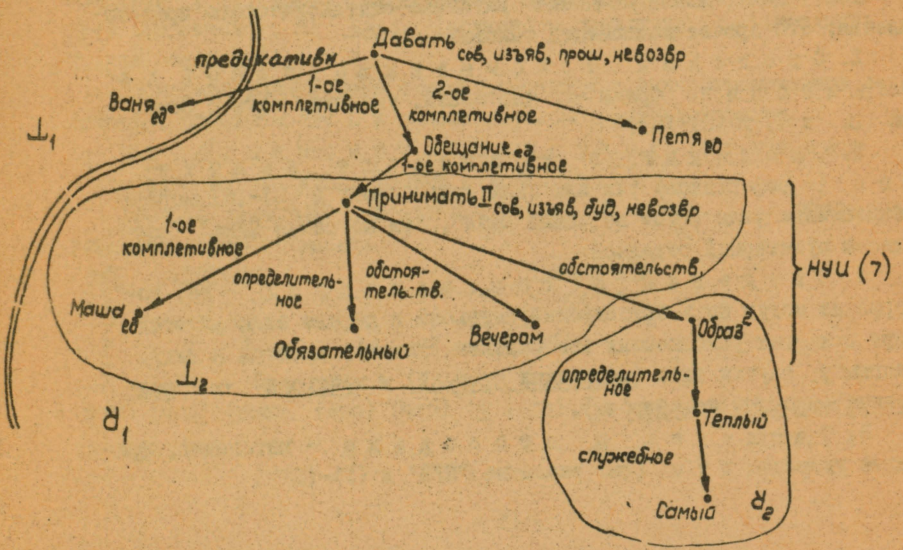
3. Сведения о тождестве денотатов (= референтов) тех или иных групп слов; в (2)-(5) кореференциальные узлы ("тот же самый Ваня", "та же самая Маша") соединены пунктирной стрелкой.

4. Сведения о составляющих – те, которые не могут быть естественно выражены в дереве зависимостей (при принятых условностях изображения связей); имеются в виду, например, случаи типа (медленные протоны) и нейтроны vs. медленные (протоны и нейтроны).

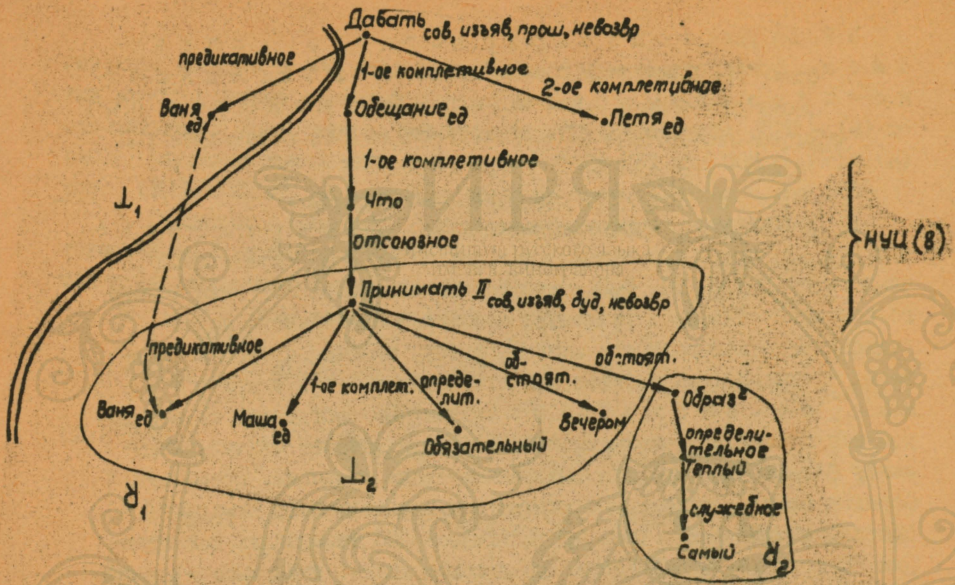
5. Сведения о просодике – интонация, фразовые ударения и т.п. (см. указание "НУИ" в (2)-(5)).



Ваня твердо обещал Пете, что вечером Маша найдет у него самый теплый прием.

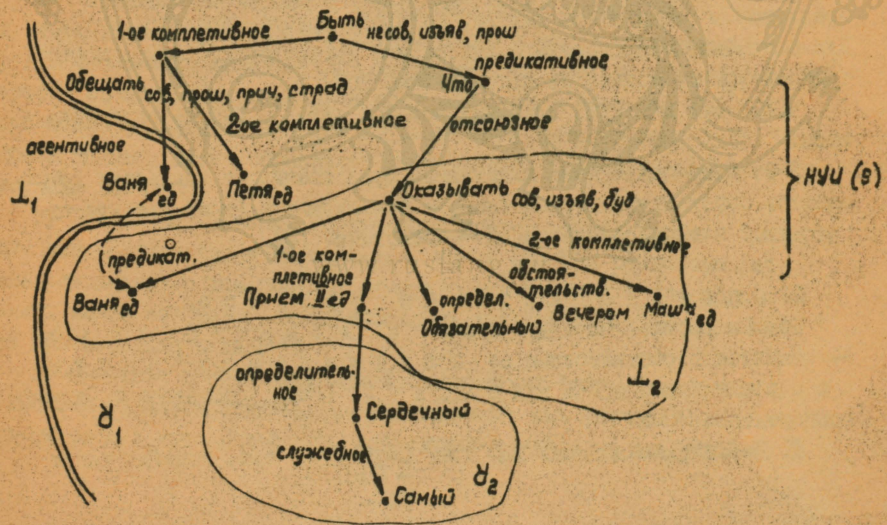


Ваня дал Пете обещание вечером обязательно принять Машу самым теплым образом.



Ваня дал Пете обещание, что вечером он обязательно примет Машу самым теплым образом.

66823



Всей было обещано Пете, что вечером он обязательно окажет Маше самый сердечный прием.

БИБЛИОТЕКА
ИН-ТА РУССКОГО ЯЗЫКА
ИН-ТА ЯЗЫКОЗНАНИЯ АН СССР

II В. На уровне поверхностного синтаксиса фразе ставится в соответствие ее поверхностно-синтаксическое представление (ПСП); см.(6) – ПСП, отвечающее ГСП (3), (7)–(8) – ПСП, отвечающие ГСП (4), (9) – ПСП для ГСП (5).

ПСП фразы также состоит из пяти компонентов.

1. Центральным компонентом является поверхностно-синтаксическая структура фразы (ПСС), аналогично ГСС в ГСП. ПСС – это дерево зависимостей, узлы которого помечены символами реальных лексем данного языка (здесь представляются все лексемы фразы, включая служебные), а ветви – символами поверхностно-синтаксических отношений (ПСО). Каждое ПСО соответствует некоторой конкретной синтаксической конструкции данного языка – комплексу формальных средств (морфологические признаки, просодемы, порядок слов), связанных с выражением определенного комплекса смысловых отношений. Так, предикативное ПСО изображает конструкцию "подлежащее + сказуемое" (в русском языке: сказуемое – в лице и числе подлежащего / в непрош. времени, в числе и роде подлежащего/ прош.времени), отпредложное ПСО – конструкцию "предлог + существительное" (в русском языке: существительное обязательно следует за предлогом и стоит в том падеже, которого требует предлог) и т.п. Для описания поверхностного синтаксиса естественного языка может требоваться несколько десятков разных ПСО.

В : как и в ГСС, линейный порядок узлов в ПСС не определяется.

2–5. Сведения о коммуникативной организации, о тождестве денотатов определенных групп слов, о составляющих (там, где это необходимо) и о просодике фразы – аналогично ГСП.

III. Морфологический уровень.

IIIА. На уровне глубинной морфологии фразе ставится в соответствие ее глубинно-морфологическое представление (ГМП), состоящее из двух компонентов:

1. Последовательность ГМП всех словоформ, т.е. линейно упоря-

доченное множество ГМП словоформ фразы. ГМП словоформы – это имя лексемы вместе с полной морфологической характеристикой, однозначно (с точностью до свободного морфологического варьирования) задающие конкретную словоформу: СТОЛБ$_{мн,род}$ (столбов), СВОЙСТВЕННЫЙ$_{ед,муж,кр}$ (свойственен/свойствен) и т.п.

2. Сведения о просодике фразы – ритмические группы, их интонационные контуры, эмфатические выделения и т.п. (При ориентации на письменный текст указываются сведения о пунктуации.)

См. (10) – ГМП фразы, отвечающее ПСП (6):

(10)
- ВАН$^{3/4}_{ед,им}$ | ТВЕРДЫЙ$_{ср,ед,кр}$ ОБЕЩАТЬ$_{сов,изъяв,прош,невозвр,ед,муж}$
- ПЕТЯ$_{ед,дат}$ ‖ ЧТО МАША$^{3/4}_{ед,им}$ | НАХОДИТЬ$_{сов,изъяв,буд,невозвр,3\ ед}$
- У ОН$_{ед,род}$ ВЕЧЕРОМ$^{3/4}$ | САМЫЙ$_{муж,ед,вин}$ ТЕПЛЫЙ$_{муж,ед,вин}$
- ПРИЕМ П$^{1/2}_{ед,вин}$ ‖

Краткий комментарий к ГМП (10).

1) Вертикальные черточки показывают паузы разной длительности, т.е. границы ритмических групп. Цифра в конце ритмической группы обозначает определенный интонационный контур, а косая черта – свободное варьирование контуров (таким образом, ГМП (10) изображает в действительности 10 по-разному интонируемых ГМП). Жирные знаки акцента обозначают эмфатическое выделение соответствующих слов.

2) В отличие от обоих "древесных" представлений (фраз), где линейный порядок узлов не задается, здесь все ГМП словоформы расставлены именно в том порядке, какой соответствующие словоформы должны иметь во фразе: линеаризация дерева фразы выполняется как раз при переходе к ее ГМП.

3) Сведения о коммуникативной организации фразы (компонент 2 в ГСП и ПСП) здесь отдельно не указываются: они выражаются порядком слов и просодикой (ср. эмфатическое ударение на САМЫЙ и

ТЕПЛЫЙ, передающие $\mathcal{A}_2$).

4) Сведения о тождестве денотатов (компонент 3 в ГСП и ПСП) здесь также отдельно не указываются: они выражены выбором местоимений-заменителей, ср. ОН вместо второго вхождения лексемы ВАНЯ.

III В. На уровне поверхностной морфологии фразе ставится в соответствие ее поверхностно-морфологическое представление (ПМП), состоящее из двух компонентов, как и ГМП:

1. Последовательность ПМП словоформ; ПМП словоформы – это цепочка морфем и символов морфологических операций, задающих словоформу.

2. Сведения о просодике (пунктуации) фразы.

См. ПМП фразы (11), отвечающее ГМП (10):

(11)
{ВАНЯ}+{ЕД.ИМ}||{ТВЕРДЫЙ}+{КР.СР.ЕД}
{ОБЕЩАТЬ}+{СОВ}+{ТЕМА}+{ПРОШ}+{МУЖ.ЕД}
{ПЕТЯ}+{ЕД.ДАТ}||{ЧТО} {МАША}+{ЕД.ИМ}|
+{НАХОДИТЬ}+{СОВ}+{ИЗЪЯВ.НЕПРОШ.3 ЕД}
{У}{ОН}+{ЕД.РОД}{НЕЧЕРОМ}|{САМЫЙ}+{МУЖ.ЕД.ВИН}|||
{ТЕПЛЫЙ}+{МУЖ.ЕД.ВИН} {ПРИЕМ}+{ЕД.ВИН}

IV. Фонологический уровень: фонологическое представление (ФП) фразы, с указанием всех просодем; см. ФП (12), отвечающее ПМП (11):

(12)
/vân'a^{3} | tv'órdo ob'eščál p'ét;e^{3} || čto mášа^{4} najd'ót u
n'ovo^{4} v'éčorom | sámij t'óplij pr'ijám^{1} |||/

V. Фонетический/графический уровень: либо фонетическая транскрипция фразы, с указанием всех просодий; либо запись фразы в принятой орфографии с соответствующей пунктуацией, см. (13) и (14):

(13) [ván'ъ tv'órdъ ʌb'išál p'ét'i‖ što maš̆ənaid'ótun'ivo|
v'éč̆ ir̆m šámъi t'óplъi pr'ijóm‖|]

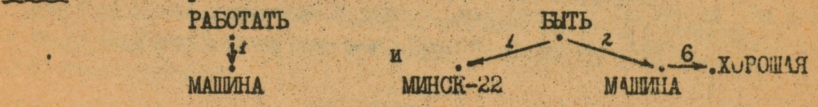
(Здесь в словах самый и теплый показана эмфатическая долгота начальных согласных.)

(14) Ваня твердо обещал Пете, что вечером Маша найдет у него самый теплый прием.

Теперь, опираясь на введенные понятия и примеры (1)-(14), мы поясним само понятие уровня представления.

Каждый уровень n - это своего рода особый искусственный язык L_n, конструируемый для описания естественных языков в некотором определенном аспекте; а именно, L_n - это набор единиц (представляющих единицы описываемого языка) и правила их комбинирования (отражающие наблюдаемые законы сочетаемости языковых единиц). Представление уровня n можно трактовать как текст на языке L_n - построенный из его единиц и по его правилам и являющийся переводом (на L_n) описываемого конкретно-языкового высказывания.

Общеизвестно, что в естественных языках информация выражается, как правило, неявным, сложным и неоднозначным образом. Так, в русском сегмент /а/ может быть отдельным словом - союзом а, отрицательным префиксом (анормальный), флексией им. пад. ед. числа (стен-а, дяд'-а), флексией род./вин. пад. ед. числа (стол-а, слон-а), флексией им./вин. пад. мн. числа (остров-а, учител'-а), темой глагола (стон-а-ть), суффиксом деепричастия (привед'-а) и т.д. На морфологическом уровне эта неоднозначность снимается: -а в стена записывается как "ед,им", -а в острова - как "мн,им"/"мн,вин",... Однако такие записи, как "ед,им", ..., в свою очередь тоже неоднозначны: слово в им. падеже может быть подлежащим, именной частью сказуемого, приложением (в системе человек-машина). Эта следующая неоднозначность снимается на синтаксическом уровне; так, машина в ГСС фраз Машина работает и в "Минск-22" - это хорошая машина занимает разные места:



Но и здесь не достигается полная однозначность: зависимый член одного и того же ГСО может играть разные семантические роли; так, при ГСО № 1: Иван любит - испытывает чувство, Иван воспитывает... - занят целесообразной деятельностью, Иван воспитывается - подвергается действиям, Иван толстеет - характеризуется изменением свойства, Иван падает - перемещается и т.п. Только на семантическом уровне отношение между Иваном и любовью, Иваном и воспитанием, Иваном и телесной толщиной должно выражаться максимально явным и однозначным образом.

Тем самым, чем "глубиннее" уровень, тем большую эксплицитность в записи содержания высказываний он обеспечивает. Семантический язык, т.е. язык семантического уровня, должен гарантировать максимальную эксплицитность и однозначность изображения всей смысловой информации (независимо от того, как она представляется в соответствующих текстах). Язык глубинно-синтаксических представлений позволяет эксплицитно и однозначно выразить обобщенные синтаксические и лексико-сочетаемостные [но не семантические!] связи между полнозначными лексическими единицами; тем самым, на этом языке выразим ряд формальных особенностей текста (членение на фразы и на лексические единицы, общая схема строения фразы). Еще больше о реальном тексте (но зато меньше о его смысле) сообщает в явном виде запись на языке ПСП: непосредственно виден полный лексический состав и реальные синтаксические конструкции текста, изображенные при этом абсолютно эксплицитно и однозначно (именованными стрелками). На морфологических уровнях вся морфологическая информация выражается эксплицитно, однако синтаксические отношения в явном виде не показаны: они "запрятаны" в значения тех или иных морфологических переменных, в линейный порядок ГМП и т.д.

В то же время язык любого уровня представления должен обеспечивать сохранение в с е й смысловой, содержательной информации, передаваемой высказыванием. Это значит, что от ПСП и ГСП, ГМП и ГМ фразы требуется задание фразы (взятой в о д н о м значении) с точностью до полной синонимии. В то же время на каждом уровне, более глубинном, чем сам текст, омонимия, во-

обще говоря, должна быть не больше, чем в тексте (т.е. при движении к смыслу омонимия должна только уменьшаться, полностью снимаясь в Сем.П.).

NB : намеченный здесь состав уровней не является окончательным. Кроме указанных основных уровней, при установлении соответствий "смыслы ⟷ тексты" оказывается целесообразным использовать и другие уровни, например, лежащие между поверхностно-синтаксическим и глубинно-морфологическим ("уровень членов предложения"), между поверхностно-морфологическим и фонологическим (поморфное представление словоформ) и т.д.

РАЗДЕЛ II.

Перейдем теперь к обзору основных компонентов модели "Смысл ⟷ Текст". Этих компонентов - шесть:

I. С е м а н т и ч е с к и й, устанавливающий соответствие между любым СемП и всеми альтернативными последовательностями ГСП фраз, которые образуют синонимичные высказывания, имеющие именно данное СемП.

II. С и н т а к с и ч е с к и й, устанавливающий соответствие между некоторым ГСП и всеми отвечающими этому ГСП альтернативными ГМП фраз. Компонент II состоит из двух подкомпонентов:

II А. Г л у б и н н о - с и н т а к с и ч е с к и й: от ГСП - ко всем отвечающим ему альтернативным ПСП.

II В. П о в е р х н о с т н о - с и н т а к с и ч е с к и й: от ПСП - ко всем отвечающим ему альтернативным ГМП фраз

III. М о р ф о л о г и ч е с к и й, устанавливающий соответствие между ГМП словоформы и ее фонологическим представлением. Здесь также имеются два подкомпонента:

III А. Г л у б и н н о - м о р ф о л о г и ч е с к и й: от ГМП словоформы к ее ПМП.

III В. П о в е р х н о с т н о - м о р ф о л о г и ч е с к и й: от ПМП словоформы - к ее фонологическому представлению.

IV. Ф о н о л о г и ч е с к и й, устанавливающий соответствие между фонологическим представлением словоформы и ее фонетической транскрипцией.

У. Графическо-орфографический, устанавливающий соответствие между фонологическим представлением воформы и ее орфографической записью (сюда же относится пунктуация).

УI. Фонетический, устанавливающий соответствие между фонетической транскрипцией и реальным звучанием.

NB: место и номер графическо-орфографического компонента зависят от характера орфографии, принятой в данном языке. Если орфография фонемна (т.е. буквами изображаются в основном фонемы), как в русском языке, то он может работать вслед за морфологическим компонентом, тогда он как бы параллелен фонологическому компоненту. Если же орфография фонетична (буквами изображаются звуки речи), как в белорусском языке, то этот компонент должен идти за фонологическим; тогда его входом является не фонологическое, а фонетическое представление.

I. Семантический компонент модели "Смысл ⟷ Текст". Чтобы полностью охарактеризовать его, требуется рассмотреть три вопроса:

(а) Семантическое представление высказываний, т.е. семантический язык.

(б) Глубинно-синтаксическое представление фраз, т.е. строение ГСС, прочие компоненты ГСП, лексические функции и т.д.

(в) Правила перехода СемП ⟷ ГСП: их типы и выполняемые ими операции.

NB: и вообще, для описания любого n-го компонента нашей модели требуется охарактеризовать уровень представления n, уровень представления n + 1 и правила перехода между ними.

Поскольку относительно (а) и (б) кое-что было сказано выше, здесь мы сообщим нечто (столь же немногое) только о (в).

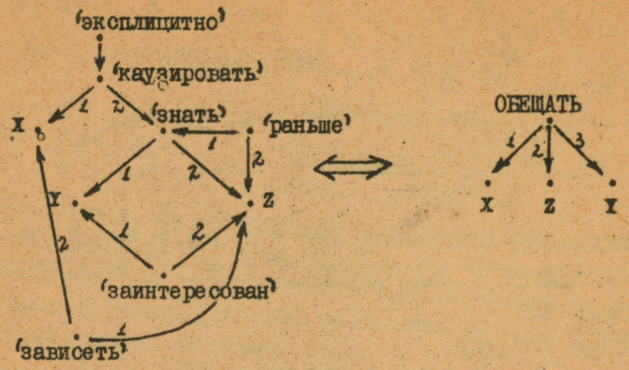
При движении от смысла к тексту семантический компонент выполняет следующие операции (при обратном движении все эти операции выполняются "наоборот")[1]:

1. Расчленяет СемП на куски, отвечающие каждый одной фразе.

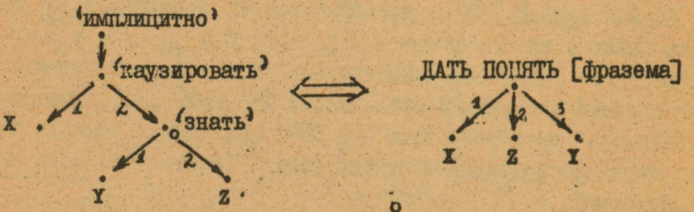
[1] Эти операции вовсе не должны выполняться обязательно в том порядке, в котором они здесь перечислены; вообще, вопрос о последовательности преобразований нами здесь сознательно даже не ставится.

2. Подбирает обобщенные лексемы посредством правил типа:

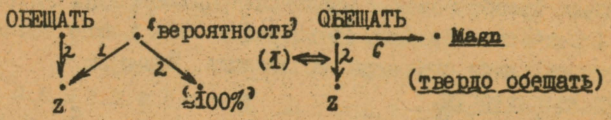
(15)

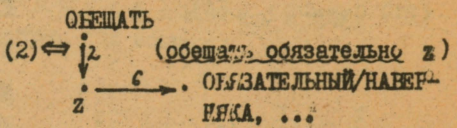


(16)



(17)





3. Подбирает семантические (т.е. синтаксически не обусловленные) морфологические характеристики лексических единиц посредством правил типа (18)-(20):



(18) $X \Leftrightarrow .X_{прош}$ | X – глагол

(19) X .◄——. 'один/однократно' $\Leftrightarrow .X_{сов}$ | X – глагол

(20) X .◄——. 'один/однократно' $\Leftrightarrow .X_{ед}$ | X – существительное

Н : любопытно отметить соответствие между сов. видом и ед. числом, ср. ударил : удар= ударял : удари или же прыгнул: прыгал = прыжок: прыжку .

4. Строит ГСС фразы.

5. Вырабатывает остальные компоненты ГСП.

6. Для каждой ГСС специальный раздел семантического компонента – так называемая с и с т е м а п е р и ф р а з и р о в а н и я – строит все такие синонимичные ей ГСС, что эта синонимия может быть описана в терминах лексических функций. Иначе говоря, задается "алгебра" преобразований над ГСС, содержащих символы ЛФ. Благодаря этим преобразованиям учитывается значительная доля синонимии высказываний.

Данные преобразования задаются правилами двух классов.

– Лексические правила представляют собой либо смысловые эквивалентности, либо смысловые импликации. Примеры:

Эквивалентности.

(21) $C \overset{X}{\Leftrightarrow} \overset{Y}{Conv_{21}}(C)$ — Множество A содержит также точку x. $\overset{C}{\Leftrightarrow}$ Множеству A принадлежит также точка x. ($Conv_{21}(C)$)

(22) $C \overset{X}{\Leftrightarrow} \overset{Y}{Adv_{1B}}(C)$ — Он поспешил выйти. $\overset{C}{\Leftrightarrow}$ Он поспешно вышел. ($Adv_{1B}(C)$)

(23) $C \overset{X}{\Leftrightarrow} \overset{Y}{S_0}(C) + \overset{Z}{Oper_1}(S_0(C))$ — Он борется. $\overset{C}{\Leftrightarrow}$ Он ведет борьбу. ($Oper_1\ S_0(C)$)

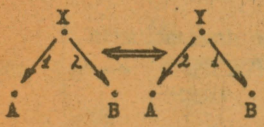
(24) $Real_2(C) \overset{X}{\Leftrightarrow} \overset{Y}{Adv_{1L}}(Real_2(C))$ — Он последовал ее совету уехать. $\overset{C}{\Leftrightarrow}$ Он уехал по ее совету. ($Adv_{1B}Real_2$)

Импликации[1]

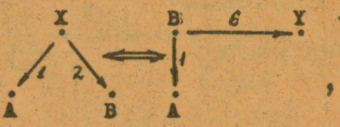
(25) $PerfCaus(X) \Rightarrow PerfIncep(X)$ Петр $\overset{PerfCausFunc_0(C)}{\overset{C}{\text{запустил}}}$ мотор. $\Rightarrow$ Мотор $\overset{PerfIncepFunc_0(C)}{\overset{C}{\text{заработал}}}$.

(26) $PerfIncep(X) \Rightarrow X$ Мотор $\overset{PerfIncepFunc_C(C)}{\overset{C}{\text{заработал}}}$. $\Rightarrow$ Мотор $\overset{Func_0(C)}{\overset{C}{\text{работает}}}$.

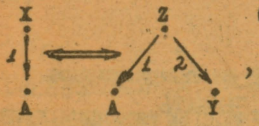
– Синтаксические правила описывают преобразования деревьев зависимостей, указывая, какие перестройки ГСС нужны в случае применения того или иного лексического правила. Так, для обслуживания лексических правил (21)–(24) необходимы, соответственно, синтаксические правила следующих типов:

(27) 

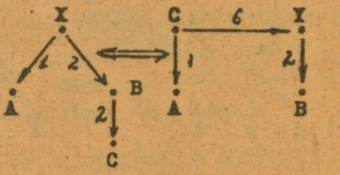
,

(28) 

,

(29) 

,

(30) 

II. С и н т а к с и ч е с к и й к о м п о н е н т.

II, А. Говоря о г л у б и н н о – с и н т а к с и ч е с к о м компоненте, нужно рассмотреть опять-таки три вопроса:

(а) ГСП,

(б) ПСП,

(в) Правила перехода ГСП $\Longleftrightarrow$ ПСП.

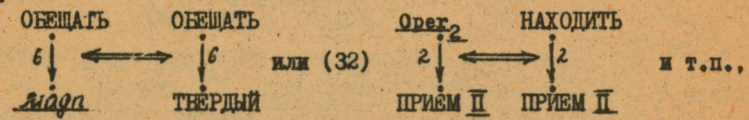
Как и выше, мы коснемся здесь только правил перехода, состоящего из двух преобразований.

---

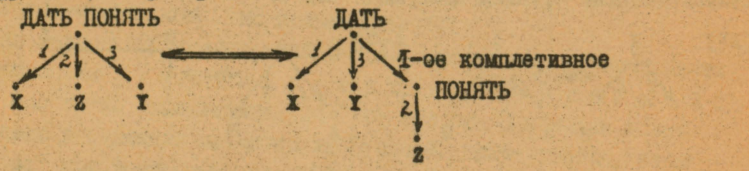
[1] Вообще говоря, смысловые импликации описывают не синонимические преобразования; однако в ряде специфических контекстов они могут использоваться и в качестве синонимических преобразований, ср., например, (25) в случаях типа По нашей улице пустили троллейбус. $\Longleftrightarrow$ По нашей улице пошел троллейбус.

1. Преобразование ГСС⟷ПСС. Соответствующие правила (аналогично правилам системы перифразирования) распадаются на два класса:

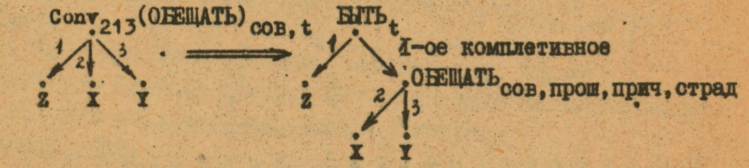
- Лексические правила преобразуют узлы деревьев. Они "вычисляют" значения лексических функций (при данном аргументе), например:



а также разворачивают в поверхностно-синтаксические поддеревья узлы-фрземы, например:



и узлы аналитических форм:

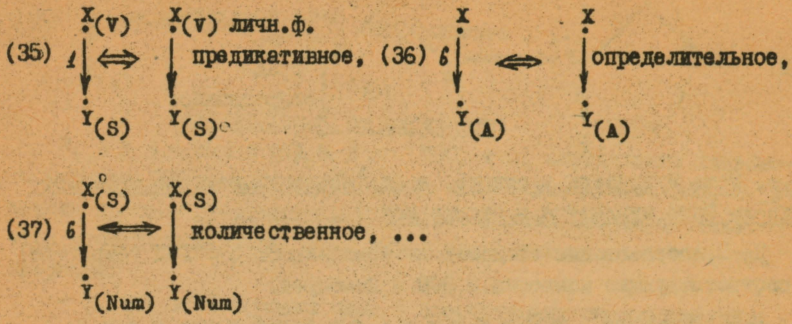


(т.е. быть обещанным - аналитический пассив к обещать; t - время глагола) и т.п.

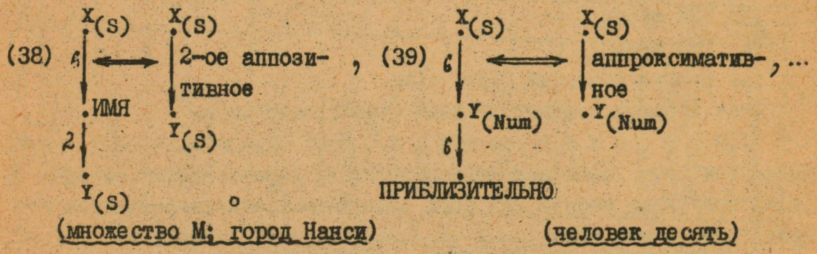
Все это делается на основе информации, которая содержится в специальном словаре, составляющем существенную часть модели "Смысл ⟷ Текст".

Синтаксические правила преобразуют само дерево, выполняя операции следующих четырех типов:

- Замена ГСО на ПСО, например:

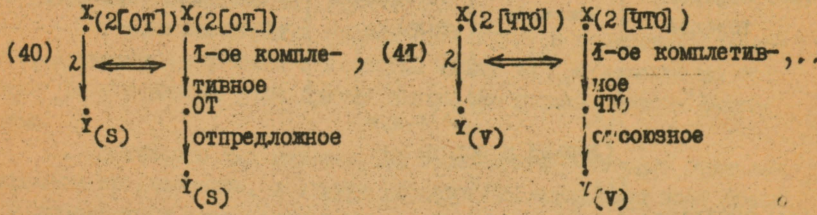


- Замена глубинного узла на ПСО, например:



(множество М; город Нанси) (человек десять)

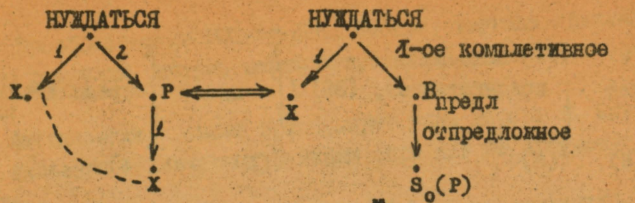
- Замена ГСО на поверхностный узел, например:



(запись вида $X_{(1[z])}$ означает "Лексема, 1-ая валентность которой оформляется в ПСС служебным словом-предлогом или союзом - Z"; эти сведения хранятся в только что упомянутом словаре).

- Элиминация узлов ГСС, участвующих в анафорических связях и не подлежащих выражению в тексте, например:

(42)



(т.е. Маша нуждается в отдыхе, а не *Маша нуждается в своем отдыхе ⟺ Маша нуждается в отдыхе Маши).

2. Преобразование сведений из компонентов 2-5 ГСП (стр.11-12) в соответствующие компоненты ПСП и наоборот.

II В. Описание п о в е р х н о с т н о - с и н т а к с и ч е с к о г о компонента требует рассмотрения:

(а) ПСП фразы,

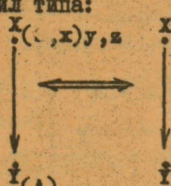
(б) ГМП фразы,

(в) правил перехода ПСП⟺ГМП.

Этот переход состоит из пяти преобразований.

1. М о р ф о л о г и з а ц и я ПСС: определение всех синтаксически обусловленных морфологических характеристик лексем (т.е. получение набора ГМП всех словоформ фразы); выполняется посредством правил типа:

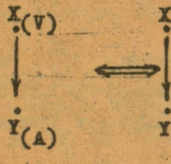
(43) определительное



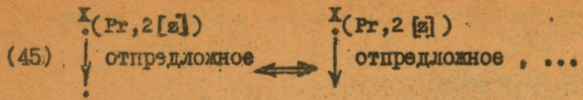
X не имеет при себе числительного
(x - род, y - число, z - падеж)

(ловкий поворот)

(44) определительное



(ловко повернулся)



(45)

(к столбцам от плоскости)

2. Линеаризация ПСС, т.е. определение линейного порядка ГМП словоформ ("порядка слов") для кусков ПСС, отвечающих простым предложениям в составе фразы. Это осуществляется посредством правил трех типов, выполняющих следующие три операции:

- Определение порядка слов в элементарных синтаксических группах в соответствии с шаблонами типа

1	2	3	4	5	6	7	8	9	10	11
сочин. союз	отрицат. или огран. частица	предлог	квантор. слово	указат. А	количеств. слово	притяжат. А	порядковое А		одно (= элементарная) группа	приложение

и не из всех этих трех наших вторых в ных формул 2

- Определение порядка элементарных групп внутри приведенных (= полных) синтаксических групп.

- Определение порядка приведенных групп в простом предложении с учетом их состава и свойств, членения предложения на тему - рему и целого ряда других сложно взаимодействующих факторов.

3. Соединение тех цепочек ГМП (словоформ) которые отвечают простым предложениям, в единую цепочку ГМП словоформ всей фразы (с выполнением всех необходимых изменений в соединяемых цепочках).

4. Прономинализация: введение в цепочку ГМП словоформ фразы местоимений.

5. Выполнение над цепочкой ГМП обязательных и факультативных эллипсисов.

Ш. Морфологический компонент.

Ш А. Глубинно-морфологический компонент определяется:

(а) множеством ГМП словоформ,

(б) множеством морфемных цепочек,

(в) правилами перехода ГМП $\Longleftrightarrow$ цепочка морфем.

Эти правила имеют вид:

(46) $T_{(s)} \Longleftrightarrow \{T\}_{(s;o)}$

(47) $y,z \Longleftrightarrow \{YZ\}_{(1)}$

(48) $T_{(v;тем)} \Leftrightarrow \{T_{v;тем;o}\}\{ТЕМА\}_{(3)}$

наст, $u,y \Rightarrow \{$НЕПРОШ. $u,y\}_{(4)}$

Здесь: Т-основа; y,z -число, падеж; u -лицо; цифра в скобках при морфеме - "порядок морфемы", т.е. ее относительная позиция в словоформе; ТЕМА -морфема, алломорфами которой являются так называемые тематические элементы в глаголе (крич-а-ть и т.п.)

Например, в результате применения правил этого компонента

ГМП = СБЛИЖАТЬСЯ$_{\text{сов,прош,прич,дейст,мн,дат}}$

получит вид

$\{$СГЛИЖ/ТЬСЯ$\}_{(0)}$ + $\{$СОВ$\}_{(1)}$ + $\{$ТЕМА$\}_{(3)}$ + $\{$ПРИЧ.ДЕЙСТ.ПРОШ$\}_{(4)}$ +

сблиз + и + вш +

+ $\{$МН.ДАТ$\}_{(5)}$ + $\{$СЯ$\}_{(6)}$

+ ым + ся

Ш В. Поверхностно-морфологический компонент определяется:

(а) множеством морфемных цепочек,

(б) множеством словоформ в фонемной транскрипции,

(в) правилами перехода "морфемная цепочка $\Longleftrightarrow$ словоформа".

Правила (в) распадаются на три группы:

- Морфемно-морфные правила типа

(49) $\{$ПРИЪМ$\} \Longleftrightarrow$ /pr'ɪjom/$_{(P_1)}$

(50) {ТВЕРДЫЙ} ⟺ /tv'ord/$_{(p_2)}$ } Здесь и ниже p_1, p_2 ...-с и н т а к т и к а при морфах: все необходимые сведения о сочетаемости этих морф (часть речи, тип склонения, род, возможные чередования,...).

(51) {ОН} + {ЕД.ИМ} ⟺ /on/ } /on/ - м е г а м о р ф а, реализующая сразу две морфемы

(52) {ОН} ⟺ /j/ | не в им. пад | не после предлога
⟺ /n'/ | после предлога

(53) {НАХОДИТЬ} + {СОВ} ⟺ /najd/$_{(p_3)}$ | не в прош.времени } /najd/ и /našed/ и /našod/ мегаморфы, реализующие сразу по две морфемы
⟺ /našed/$_{(p_4)}$ | в прош. вр.прич.
⟺ /našod/$_{(p_5)}$ | в про.вр. не прич.

(54) {ЕД.ВИН} ⟺ Ø | либо II скл, муж,неод (ров)
| либо III скл,муж/жен (путь, ночь)
⟺ /a/ | либо II скл,од (кота)
| либо III скл,ср (время)
. . . . . .

(55) {КР.СР.ЕД} ⟺ /o/

(56) {ИЗЪЯВ. НЕПРОШ. 3 ЕД} ⟺ /ot/$_{(p_6)}$ | I спр
⟺ /it/$_{(p_7)}$ | II спр

- Правила акцентуации, выполняющие постановку ударения в словоформе в соответствии со сведениями об акцентуационном типе парадигмы; эти сведения должны быть записаны при основе словоформы (в ее синтактике):

/pr'ijom+Ø/ ⟺ /pr'ijóm+Ø/,

/najd+ot/ ⟺ /najd+ót/ [необходимое перед /ot/ смягчение выполняют правила следующей группы!] и т.д.

- Морфонемные правила, выполняющие разного рода чередования фонем, усечения и т.п.

(57) /d/ ⟹ /ž/ (/tv'órd + e/ ⟹ /tv'órž + e/)

(58) /o/ ⟹ /ó/ (/pt'ic + k + a/ ⟹ /pt'icó + k + a/)

(59) /C/ ⟹ /C'/ (/najd+ɔt/⟹/najd'+ɔt/, /tɔn+ot/⟹/tɔn'+ot/,...) } /C/ - согласная, /C'/ - ее палатализованный коррелят

(60) /t/,/d/ ⟹ Λ (/pl'ɔt+l/⟹/pl'ɔ+l/, /našɔd+l/⟹/našɔ+l/, /najd+t'l/ /naj+t'l/,...) } Λ - пустая цепочка, так что запись "X ⟹ Λ" означает "Стереть X".

(61) /Г/ ⟹ Λ (/pɔlok+a/⟹/pɔlk+a/, /ot'ec+ăm/⟹/otc+ăm/,...) } "беглые гласные"

Применение морфонемных правил контролируется как синтактиками основ (указаниями о том, в какой основе и при каких условиях какое чередование возможно/обязательно и т.п.), так и синтактиками самих морфонемных правил (указаниями о том, в каких условиях и к чему данное правило может/должно быть применено).

Среди морфонемных правил имеются также правила, обеспечивающие выполнение чередований, усечений и т.п. на стыках словоформ - т.е. в масштабе фразы.

IV. Фонологический компонент:

1. Устанавливает соответствие между фонологическим представлением словоформы и ее фонетической транскрипцией, т.е. содержит правила реализации фонем их аллофонами (= звуками речи) в зависимости от позиции.

2. Устанавливает соответствие между просодической характеристикой словоформ и фразы и конкретными просодиями, отмечаемыми в фонетической транскрипции фразы.

V. Графическо-орфографический компонент:

1. Устанавливает соответствие между фонологическим представлением словоформы и ее орфографической записью, в частности - посредством правил вида:

(62) /Xja/⟺/X/я /struja/⟺струя

(63) /Xju/⟺/X/ю X-не согласная /struju/⟺струю

(64) /Xji/⟺/X/и /struji/⟺струи

(65) /jX/⟷й /X/ | X - не гласная /kraj/⟷край
(66) /С'j/⟷/Сь /j/ | /С/ - согласная /sv'in'ja/⟷свинья
(67) /С'a/⟷/С/а | /m'aso/⟷мясо

а также вида:

(68) /Ш/#⟷/Ш/ъ# | лу ↑ э ⟷ (s, жен, Ш); либо у(v) | мягкий знак после шипящих в основах (х) сущ. женского рода Ш скл. ь в глагольных формах | (мышь, читаешь, речь, спрячь)

(69) /ijе/⟷ии | /е/ ∈ {ЕД.ДАТ}, {ЕД.ПРЕ Л} (/l'in'ij+e/⟷линии)

и т.д.

2. Устанавливает соответствие между просодической характеристикой фразы и соответствующими знаками препинания. (Часть знаков препинания обусловлена синтаксически, и их постановка определяется синтаксическим компонентом.)

УI. Ф о н е т и ч е с к и й компонент: устанавливает соответствие между фонетической транскрипцией фразы и реальными физическими параметрами акустического сигнала. (Этот компонент по существу выводит нас за пределы собственно модели "Смысл ⟷ ⟷ Текст": он образует модель "Текст ⟷ Звучащая речь", о чем см. ниже.)

Таковы основные компоненты рассматриваемой модели "Смысл⟷ ⟷Текст". Отметим, что они точно соответствуют традиционно выделяемым разделам науки о языке:

I. Семантика.
II. Синтаксис.
Ш. Морфология.
IУ. Фонология.
У. Правописание.
УI. Фонетика.

Это, разумеется, неслучайно: ведь каждый компонент модели есть модель некоторого разумно выбранного фрагмента языка, а соответствующие разделы лингвистики посвящены изучению именно этих фрагментов. Такое положение вещей позволяет нам более точно формулировать задачи и содержание каждого раздела лингвистики, ссы-

лаясь на соответствующие компоненты модели "Смысл ⟷ Текст".

В заключение отметим, что модель "Смысл ⟷ Текст" должна входить в состав полной модели языкового поведения: "Действительность ⟷ Речь", в рамках которой наша модель занимает среднее место, т.е.

"Действительность ⟷ (1) Смысл ⟷ (2) Текст ⟷ (3) Звучащая речь".

Однако подмодели (1) и (3) этой полной модели лишь отчасти относятся к компетенции "чистой" лингвистики, и потому мы пока отказываемся от их рассмотрения.

x

x x

Список сокращений и некоторых терминов.

Высказывание (стр. 5, сноска).
ГМП – глубинно-морфологическое представление (стр.18).
ГСО – глубинно-синтаксическое отношение (стр. 14).
ГСП – глубинно-синтаксическое представление (стр. 10).
ГСС – глубинно-синтаксическая структура (стр.10,15).
КС – коммуникативная организация (стр. 9).
ЛФ – лексическая функция (стр.13-14).
Мегаморфа (стр. 33).
НУИ – нейтрально-утвердительная интонация (стр.11,15).
Обобщенная лексема (стр. 13).
ПМП – поверхностно-морфологическое представление (стр. 20).
Приведенная синтаксическая группа (стр.31).
ПСО – поверхностно-синтаксическое отношение (стр. 18).
ПСП – поверхностно-синтаксическое представление (стр. 18).
ПСС – поверхностно-синтаксическая структура (стр. 18)
Рема (стр. 9).
СГ – семантический граф (стр. 8).
СЕ – семантическая единица (стр. 8).
Сема (стр. 8).
СемП – семантическое представление (стр. 6).
Синтактика (стр. 33).
Тема: 1) стр. 9 ; 2) стр. 32
Уровень представления (стр. 21 и сл.).
ФП – фонологическое представление. стр.20
Фразема (стр. 13).
Элементарная синтаксическая группа (стр.31).

Литература.

[1] Гладкий А.В., Мельчук И.А. Грамматики деревьев. I. Опыт формализации преобразований синтаксических структур естественного языка. – В кн.: "Информац. вопросы семиотики, лингвистики и автомат. перевода", вып. 1, 1971, 16-41.

[2] Еськова Н.А., Мельчук И.А., Санников В.З. Формальная модель русской морфологии. I. Формообразование существительных и прилагательных. М., 1971. (Предв. публикации ПГЭПЛ.)

[3] Жолковский А.К. – Статьи в сб. "Машинный перевод и прикладная лингвистика", вып. 8, 1964.

[4] Жолковский А.К. О глубинном и поверхностном синтаксисе (на материале языка сомали). – "Изв. АН СССР. Сер. лит. и яз., 29, 1970, № 5, 427-439.

[5] Жолковский А.К. Синтаксис сомали (глубинные и поверхностные структуры). М., 1971.

[6] Жолковский А.К., Мельчук И.А. О семантическом синтезе. – "Проблемы кибернетики", вып. 19, 1967, 177-238.

[7] Жолковский А.К., Мельчук И.А. К построению действующей модели языка "Смысл —— Текст". – "Маш. перевод и прикл. лингвистика", вып. 11, 1969, 5-35.

[8] Кулагина О.С., Мельчук И.А., Эрастов К.О. Об одной возможной системе машинного перевода. М., 1971. (Предв. публикации ПГЭПЛ, вып. 21).

[9] Мельчук И.А. Об автоматическом морфологическом синтезе (на материале испанского языка). – НТИ, 1965, № 4, 35-42.

[10] Мельчук И.А. Порядок слов при автоматическом синтезе русского текста. – НТИ, 1965, № 12, 36-44.

[11] Мельчук И.А. К построению действующей модели языка. – В кн.: "Проблемы языкознания", М., 1967, 82-89.

[12] Мельчук И.А. Модель склонения венгерских существительных. – В кн.: "Проблемы структурной лингвистики, 1967", М., 1968, 326-343.

[13] Мельчук И.А. К принципам описания означаемых (о лингвистической семантике). – В кн.: "Язык и человек. Памяти П.С.Кузнецова", М., 1970, 201-217.

П. Словообразование и конверсия.

1. Цель настоящего сообщения – определить понятие конверсии как морфологического средства естественных языков, активно используемого, в частности, в словообразовании. При этом мы имеем в виду явления типа англ. (the) machine-gun 'пулемет' – (to) machine-gun 'обстреливать из пулемета', (the) shell 'снаряд' – (to) shell 'обстреливать из пушек, осыпать снарядами', (the) box 'рамочка' – (to) box 'обводить рамочкой, брать в рамочку' и т.п. Конверсия рассматривается здесь исключительно в синхронном аспекте (безотносительно к истории тех или иных пар соотносительных слов); при этом мы исходим из известного определения А.И.Смирницкого ("Конверсия – такой вид словопроизводства, при котором словообразовательным средством служит только парадигма слова"[1]), позволяя себе, однако, существенно обобщить введенное им понятие и уточнить его формулировки.

2. Сравним два ряда пар:

(а) (to) tramp 'бродяжничать' – (the) tramp 'бродяга' = 'тот, кто бродяжничает'

(to) swot 'зубрить' – (the) swot 'зубрила' = 'тот, кто зубрит',

(to) gossip 'сплетничать' – (the) gossip 'сплетник' = 'тот, кто сплетничает' и т.п.;

(б) (to) sing 'петь' – (the) sing=er 'певец' = 'тот, кто поет',

(to) smoke 'курить' – (the) smok=er 'курильщик' = 'тот, кто курит',

(to) hold 'держать' – (the) hold=er 'держатель' = 'тот, кто держит' и т.п.

Правые члены и в (а), и в (б) семантически сложнее левых на один и тот же смысл 'тот, кто...'; в (б) этот смысл передается суффиксом =er ; мы хотим называть конверсией именно

[1] В связи с конспективным характером настоящего изложения мы отказываемся от литературных ссылок и какой бы то ни было полемики.

то, посредством чего передается этот же самый смысл в правых членах пар (а). Таким образом, конверсия будет у нас н о с и т е л е м з н а ч е н и я, в каком-то отношении подобным аффиксу (в нашем примере – суффиксу =er).

3. Спрашивается: что можно счесть носителем смысла 'тот, кто...' в слове типа (the) tramp? Как кажется, здесь представляются четыре логических возможности:

(i) Считать, что этот смысл передается самой основой, т.е усматривать здесь регулярную полисемию типа рус. слива = 1) 'плод X', 2) 'дерево, на котором растут плоды X'. Тогда: tramp = 1) 'бродяжничать', 2) 'тот, кто бродяжничает'; swot = 1) 'зубрить', 2) 'тот, кто зубрит'; и т.п.

(ii) Считать, что этот смысл передается нулевым суффиксом. Тогда: tramp 'бродяжничать' – tramp=Ø 'тот, кто бродяжничает'.

(iii) Считать, что этот смысл передается каждым из "субстантивных" аффиксов и служебных слов, соединяющихся с tramp. Тогда: the = артикль опред. + 'тот, кто...', a = артикль неопр. + 'тот, кто...', =s = мн. ч. + 'тот, кто...', of = 'X + + тот, кто...' ('X' здесь – значение предлога of) и т.п.

(iv) Считать, что этот смысл передается изменением грамматической сочетаемости глагольной основы tramp – самым тем фактом, что "исконно" глагольная основа употребляется как субстантивная: с субстантивными окончаниями и в субстантивных синтагмах.

Первый способ описания форм типа (the) tramp не принимается нами из-за того, что при нем вообще не выделяется никакое словообразовательное средство и весьма специфическое морфолого-словообразовательное явление полностью сводится к обычной словарной полисемии. (Заметим, что в отличие от русской полисемии типа слива, где семантическое различие не сопровождается никаким грамматическим различием, в английском различие на 'тот, кто...' в парах типа (а) сопровождается всегда одними и теми же грамматическими различиями.)

Второй способ также лишает рассматриваемое явление его специфики, сводя конверсию к аффиксации (а именно – к нулевой). Мы же хотим определить конверсию именно как нечто специфичное, имея в виду расширить арсенал средств лингвистического описания, а потому не принимаем здесь этот способ. № : однако нулевая аффиксация эквивалентна конверсии (как мы определим ее ниже)! Это значит, что любые факты, которые могут быть адекватно описаны в терминах конверсии, могут быть столь же адекватно описаны и с помощью нулевого аффикса (ср. известную статью В.В.Лопатина). Мы не будем рассматривать здесь вопрос о критериях выбора между этими двумя равновозможными описаниями.

Третий способ очевидным образом нелеп.

Предлагается остановиться на четвертом способе описания, т.е. называть конверсией несущее смысл регулярное изменение грамматической сочетаемости.

4. Прежде чем уточнить сущность этого последнего способа, нам нужно кратко охарактеризовать важное в данной связи понятие языкового знака. Для нас языковой знак – это упорядоченная тройка вида

$$\underset{\sim}{X} = \langle \underset{\sim}{X}, \text{'X'}, \Sigma_X \rangle ,$$

где $\underset{\sim}{X}$ – означающее (то, что прямо или косвенно воспринимается), 'X' – означаемое (передаваемая знаком информация), а Σ_X – синтактика (сведения о правилах комбинирования данного знака с другими знаками). Несколько кратких пояснений:

– Мы расширяем здесь представление Ф. де Соссюра о языковом знаке как о двусторонней сущности; знак не менее, чем трехсторонен (а может быть, и четырехсторонен, включая еще и прагматику – в смысле Ч.Морриса).

– Мы расширяем традиционное представление об означающем как о звуковой цепочке; означающими признаются любые носители смысла, в частности – операции (такие, например, как значащие чередования или повторы, а также, как мы увидим, и некоторые другие).

– В синтактику помещаются только те сведения о сочетаемости знака, которые диктуются языком, но не зна-

ниями о действительности и т.п. Так, в синтактике не должен отмечаться тот факт, что слива-плод сочетается с глаголами есть, варить, покупать, а слива-дерево - с глаголами спилить или поливать: это полностью определяется пониманием реалий и не зависит от русского языка.

5. Опираясь на указанное в п. 4 понятие языкового знака, мы попытаемся построить дедуктивную систему морфологических средств естественных языков.

Пусть имеется знак $\underline{A} = \langle \underline{A}, \text{`A'}, \Sigma_A \rangle$ и означаемое `B', которое надо "прибавить" к означаемому `A'. При этом нас интересуют только м о р ф о л о г и ч е с к и е  с р е д с т в а, т.е. то, что как-то в ы р а ж а е т смысл и притом обязательно в п р е д е л а х  с л о в о ф о р м ы. Поэтому результирующий знак $\underline{AB}$ с означаемым `A + B' должен быть с л о в о ф о р м о й (или частью словоформы) и отличаться от $\underline{A}$ н а б л ю д а е м ы м о б р а з о м (отличие $\underline{AB}$ от $\underline{A}$ может усматриваться или отдельно по самому $\underline{AB}$, или по употреблению $\underline{AB}$ в тексте). Мы видим только два класса способов выполнить эти требования, т.е. два основных класса морфологических средств:

I. Прибавить к $\underline{A}$ другой отдельный знак $\underline{B}$, имеющий в качестве означающего некоторую сущность: либо цепочку фонем, либо просодему (тон, ударение); важный частный случай - нулевое означающее. В зависимости от того, чем является $\underline{B}$, мы имеем IA) а ф ф и к с а ц и ю или IB) с л о ж е н и е (= корне-, осново- и словосложение).

II. Выполнить над $\underline{A}$, а точнее - над его означающим или синтактикой некоторую операцию $\underline{O}$. В зависимости от $\underline{O}$ мы различаем здесь два подкласса:

П.I. М о д и ф и к а ц и я: если $\underline{O}$ - операция изменения означающего (разновидности: П.1.А - значащее чередование, или апофония; П. 1.В - повтор).

П.2. К о н в е р с и я - операция изменения синтактики.

Других морфологических способов в рамках предлагаемого подхода быть не может. Таким образом, конверсия занимает дедуктив-

но предсказываемую клетку в полной системе морфологических средств естественных языков.

6. Определение. Конверсией называется языковой знак, означающим которого является правило изменения синтактики других языковых знаков.

Примеры конверсий (в английском языке):

1) $\underline{К}_1$ = $\langle S \Rightarrow V$, слаб ; 'подвергать воздействию...' ; $С_1\rangle$, где $С_1$ – условия (тип и знание основ), к которым эта конверсия применима (the machine-gun – to machine-gun);
2) $\underline{К}_2$ = $\langle V \Rightarrow S$, x; 'тот, кто...' ; $С_2\rangle$, где x – морфологический тип результирующего S (to swot – a swot);
3) $\underline{К}_3$ = $\langle V \Rightarrow S$, x; 'одна порция...' ; $С_3\rangle$ (to kiss – a kiss, to look – a look, ...).

Конверсия аналогична аффиксу, а применение конверсии – присоединению аффикса.

7. Типы конверсий.

– По характеру выражаемого значения: словоизменительные/словообразовательные.

– По характеру изменяемой сочетаемости: внутренние (= внутрисловные: изменяется только сочетаемость с аффиксами)/внешние (изменяется сочетаемость с другими словоформами).

– По характеру изменяемых компонентов синтактики: парадигматические (≈ аффиксальные), категориальные (переход в другую часть речи) и рекционные (смена согласования или управления).

8. В терминах конверсий удается естественным образом описывать различные типы так называемого "морфолого-синтаксического" и "безаффиксного" словообразования в русском языке. Эти понятия вряд ли можно признать удачными: остается неясным, что же именно является с л о в о о б р а з о в а т е л ь н ы м с р е д с т в о м, т.е. чем именно передается добавляющийся смысл при "морфолого-синтаксическом" или тем более "безаффиксном" словообразовании (не говоря уже о чисто терминологической несообразности отрицательного наименования: "без аффикса"? а с чем?). Вместо этого предлагается ввести в инвентарь слово-

образовательных средств русского языка ряд конверсий, например[1]:

1) $К_1$ = ⟨ муж, П ⇒ жен, I; 'женщина [по фамилии] ...'; $С_1$ ⟩, где I и П – первое и второе склонения (Кузьмин= – Кузьмин=а, Васильев= – Васильев=а, ...);
2) $К_2$ = ⟨ муж, IУ ⇒ жен, IУ; 'женщина [по фамилии] ...'; $С_2$ ⟩, где IУ – адъективное склонение (Петровск=ий – Петровск=ая, ...);
3) $К_3$ = ⟨ муж, П ⇒ жен, Ø; 'женщина [по фамилии] ...'; $С_3$ ⟩, где Ø обозначает несклоняемость (Рабинович, ~а, ~у, ... – Рабинович, ~, ~, ...).

Синтактики $С_1$–$С_3$ определяют применимость данных конверсий только к фамилиям определенных словоизменительных типов.

4) $К_4$ = ⟨ V ⇒ S, муж, П; 'nomen actionis'; $С_4$ ⟩ (перевоз-ить – перевоз, перехват-ывать – перехват, а также отстрел, полив, обогрев, прогул, сбор, вывоз, ...); синтактика $С_4$ определяет, к каким глаголам применима эта конверсия и какие чередования должны иметь место в соответствующих основах.
5) $К_5$ = ⟨ A ⇒ S, жен, Ш; 'совокупность...'; $С_5$ ⟩ (пьянь 'совокупность пьяных', рвань 'совокупность рваного' погань, нечисть, голь, бездарь, ...);
6) $К_6$ = ⟨ A ⇒ S, жен, Ш; 'пространство/место...'; $С_6$ ⟩ (ширь, высь, даль, глушь, глубь, ...);
7) $К_7$ = ⟨ A ⇒ S, ср, IУ; '...пища'; $С_7$ ⟩ (мясное, мучное, сладкое, рыбное, кислое, свежее, ...);
8) $К_8$ = ⟨ A ⇒ S, ср, IУ; '...одежда'; $С_8$ ⟩ (переоделся в штатское; дама в голубом);
9) $К_9$ = ⟨ A ⇒ S, муж/жен, IУ; 'тот, кто...'; $С_9$ ⟩ (вахтенный, рулевой, загребной, заряжающий, дежурный, марсовой, ...);
10) $К_{10}$ = ⟨ A ⇒ S, жен, IУ; 'предприятие общественного питания, где продают...'; $С_{10}$ ⟩ (шашлычная, бутербродная, пирожковая, блинная, котлетная, чебуречная, ...);

---

[1] Значения конверсий в приводимом ниже перечне описаны достаточно условно и приблизительно, однако, как кажется, не более приблизительно, чем обычно описывают значения словообразовательных аффиксов.

11) $К_{11}$ = ⟨A ⇒ V, ej ; 'начинать быть...'; $С_{11}$⟩ (бел=еть, чёрн=еть, здоров=еть, толст=еть, бледн=еть, глуп=еть, ...);

12) $К_{12}$ = ⟨V, ej ⇒ V , и; 'каузировать ...'; $С_{12}$⟩ (белить, чернить, краснить, веселить, ...)[1].

9. Отметим сходство конверсий с аффиксами. Так, суффикс -льщик- формально должен быть изображен следующим образом: -льщик- = /l'ščik/ ; 'тот, кто...'; $С_x$ ⟩ , где синтактика $С_x$ задает те основы, к которым этот суффикс присоединим, способ его присоединения и т.п., а также содержит указание, что это именно суффикс, а не, скажем, префикс (бур-и-льщик, креп-и-льщик, провер-я-льщик, ...).

Аналогично:

-чик^1 = ⟨ /čik/ ; 'тот, кто...'; $С_y$ ⟩ (навод-чик, поднос-чик, растрат-чик, совет-чик, ...);

-щик = ⟨ /ščik/ ; 'тот, кто...'; $С_z$ ⟩ (буксиров-щик, расклейщик, расфасовщик, трамвайщик, ...).

Все эти три суффиксальные морфы объединяются в одну морфему: они полностью синонимичны и дополнительно распределены в соответствии с достаточно простыми правилами. Поэтому мы имеем:

$$\{\text{-}\check{\text{ч}}\text{ик}\} \Rightarrow /\check{s}\check{c}ik/ \mid С_z$$
$$\Rightarrow /\check{c}ik/ \mid С_y$$
$$\Rightarrow /l'\check{s}\check{c}ik/ \mid С_x$$

Точно также полностью синонимичны и дополнительно распределены конверсии $К_1$, $К_2$ и $К_3$; это позволяет нам ввести семейство конверсий, так сказать, "конверсиему", аналогичную морфеме:

$$\{К_{женск}\} \Rightarrow К_1 \mid С_1$$
$$\Rightarrow К_2 \mid С_2$$
$$\Rightarrow К_3 \mid С_3$$

В то же время суффиксы бывают омонимичны: ср. -чик^1 и -чик^2 = ⟨ /čik/ ; 'маленький...'; $С_w$ ⟩ (стульчик, карманчик, бугорчик, скандальчик, ...).

То же самое возможно и для конверсий: конверсии $К_5$ и $К_6$ омонимичны, так как у них совпадают означающие при разных означае-

---

[1] Данный список, разумеется, не исчерпывает все возможные конверсии русского языка - он преследует лишь иллюстративные цели.

мых; так же обстоит дело и с конверсиями K_7 и K_8.

Укажем еще, что в русском языке существуют и словоизменительные конверсии (это верно только при условии, что видообразование считается относящимся к словоизменению), например: $K_{несов} = \langle$ и $\Longrightarrow$ aj ; 'несов. вид'; $c_q \rangle$, где через "и $\Longrightarrow$ aj" обозначена мена тематических элементов (разделить - разделять, решить - решать, раздробить - раздроблять, ...).

10. Представляется, что использование конверсий (понимаемых в соответствии с определением пункта 6) позволяет описывать словообразование естественных языков более стандартным и одновременно более формальным образом. В частности, в терминах конверсий можно, по-видимому, гораздо точнее, полнее и систематичнее описывать разные "причудливые" словообразовательные типы русского языка, которые, как правило, остаются на периферии стандартных академических описаний, руководств и т.п.

БИБЛИОТЕКА
ИН-ТА РУССКОГО ЯЗЫКА
ИН-ТА ЯЗЫКОЗНАНИЯ АН СССР

www.ruslang.ru

Ответственный за выпуск Гончар Л.
Корректоры Хазанова И., Лебединская Н.

А05923 от 16.6.72г. 3п.л. 180экз. Зак. № 663
Тип. МГУ