

**Т.В. Батура, О.Н. Еркаева, Ф.А. Мурзин**

## **К ВОПРОСУ ОБ АНАЛИЗЕ ТЕКСТОВ НА ЕСТЕСТВЕННОМ ЯЗЫКЕ\***

### **ВВЕДЕНИЕ**

В работе рассмотрены различные понятия, конструкции и методы, применяемые в математической и классической лингвистике для анализа лексики и текстов на естественном языке.

Первые три параграфа посвящены подходу Мельчука и Аппресьяна [1, 2]. Дается формализованное описание таких понятий, как лексические функции, валентности, толково-комбинаторный словарь.

Четвертый параграф посвящен теоретико-множественным моделям Маркуса [3]. Подход Маркуса применим практически к любым языкам и позволяет на основе чисто структурного анализа определить категории рода и падежа в данном языке. Можно предположить, что посредством различных модификаций подхода Маркуса можно определить самые различные категории языка.

Далее рассматриваются словообразовательные конструкции русского языка и минимальные схемы предложений [4, 5]. Они могут быть полезны не только для морфологического и синтаксического, но и для семантического анализа.

### **1. ЛЕКСИЧЕСКИЕ ФУНКЦИИ**

Лексические функции — понятие, известное в лингвистике. Наша цель — представить лексические функции на синтаксическом уровне в виде предикатов. Позднее предполагается рассмотреть различные модели соответствующих сигнатур, и можно будет говорить об истинности данных предикатов в моделях.

Обозначим через  $L$  совокупность правильно построенных слов какого-либо естественного языка. Имеется в виду, что всевозможные словоформы,

---

\* Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант № 01-01-794) и Министерства образования РФ.

возникающие при склонениях существительных, спряжениях глаголов и др., также включены в  $L$ . Пусть  $x, y, z$  — слова или словосочетания этого языка  $L$ .

Перечень простых стандартных лексических «функций» (у нас они будут представлены в виде предикатов) приведен ниже.

1.  $Syn(x, y)$ ,  $x, y$  — синонимы.  
 $Syn_{\supset}(x, y)$ ,  $x, y$  — синонимы, но  $x$  более широкий по смыслу.
2.  $Conv(x, y)$ ,  $x, y$  — конверсивы.
3.  $Anti(x, y)$ ,  $x, y$  — антонимы.
4.  $Der(x, y)$ ,  $y$  — синтаксический дериват  $x$ , т.е.  $y$  совпадает с  $x$  по смыслу, но принадлежит к другой части речи:  
 $S_0(x, y)$ ,  $y$  — существительное, производное от  $x$  ( $x$  — не существительное).  
 $A_0(x, y)$ ,  $y$  — прилагательное, производное от  $x$  ( $x$  — не прилагательное).  
 $Adv_0(x, y)$ ,  $y$  — наречие, образованное от  $x$  ( $x$  — не наречие).  
 $V_0(x, y)$ ,  $y$  — глагол, образованный от  $x$  ( $x$  — не глагол).  
 Иначе говоря,  
 $\forall x \forall y (Der(x, y) \leftrightarrow S_0(x, y) \vee A_0(x, y) \vee Adv_0(x, y) \vee V_0(x, y))$ .
5.  $Gener(x, y)$ ,  $y$  — название понятия, родового по отношению к понятию, обозначенному  $x$  ( $x$  = «клубника»,  $y$  = «ягода»). Этот предикат зависит от лексической сочетаемости слов в данном языке: если  $x$  и  $m$  — совпадающие по смыслу слова двух разных языков, то для  $Gener(x, y)$  и  $Gener(m, n)$  соответственно  $y$  и  $n$  могут не совпадать по смыслу.

**Ситуация** — определенное лексическое отражение (в данном языке) некоторой части действительности. Ситуации, обозначаемые отдельными лексическими единицами естественных языков (лексемами), имеют, как правило, от одного до четырех «смысловых» компонентов, или **семантических актантов**, обозначаемых заглавными латинскими буквами  $A, B, C, D$ . В то же время каждой такой лексеме сопоставляются **глубинно-синтаксические актанты** — ее зависимые, соответствующие подлежащему и сильным дополнениям (в случае, если данная лексема реализуется гла-

голом-сказуемым). Глубинно-синтаксические актанты нумеруются арабскими цифрами: 1, 2, 3, 4.

6.  $S_i(x, y)$ ,  $i = 1, \dots, 4$ ,  $y$  — типовое название  $i$ -того актанта для  $x$ .
7.  $S_c(x, y)$ ,  $y$  — сирконстанта, т.е. типовое название второстепенной компоненты данной ситуации  $x$ :  
 $S_{loc}(x, y)$ ,  $y$  — типовое название места осуществления данной ситуации  $x$  (= «то, где...», например,  $x$  = «битва»,  $y$  = «поле (битвы)»);  
 $S_{instr}(x, y)$ ,  $y$  — типовое название инструмента, используемого в данной ситуации  $x$  (= «то, чем/посредством чего...»,  $x$  = «борьба»,  $y$  = «орудие (борьбы)»);  
 $S_{mod}(x, y)$ ,  $y$  — типовое название способа (манеры, характера) осуществления данной ситуации  $x$  (= «то, как...»,  $x$  = «жизнь»,  $y$  = «образ (жизни)»);  
 $S_{res}(x, y)$ ,  $y$  — типовое название результата данной ситуации (= «то, что получается»,  $x$  = «копировать»,  $y$  = «копия»);  
 Иначе говоря,  
 $\forall x \forall y (S_c(x, y) \leftrightarrow S_{loc}(x, y) \vee S_{instr}(x, y) \vee S_{mod}(x, y) \vee S_{res}(x, y) \vee \dots)$ .
8. Соотносительные предикаты  $Sign(x, y)$ ,  $y$  — типовое название одной «штуки», одного «кванта» некоторого  $x$ ;  $Mult(x, y)$ ,  $y$  — типовое название совокупности, множества.
9.  $Figur(x, y)$ ,  $y$  — метафора для  $x$  ( $x$  = «сон»,  $y$  = «объятия (сна)»).
10.  $Centr(x, y)$ ,  $y$  — типовое обозначение «центральной» части предмета или процесса.
11.  $A_i(x, y)$ ,  $i = 1, \dots, 4$ ,  $y$  — типовое определение  $i$ -го актанта по его реальной роли («такой, который...», «такой, которого...»).
12.  $Able_i(x, y)$ ,  $i = 1, \dots, 4$ ,  $y$  — типовое определение  $i$ -го актанта по его потенциальной роли в ситуации («такой, который может...», «такой, которого можно...»).
13.  $Magn_0(x, y)$  и  $Magn_i(x, y)$ ,  $i = 1, \dots, 4$ ,  $y$  обозначает «высокую степень», «интенсивность» самой ситуации  $x$  ( $Magn_0$ ) или ее  $i$ -го актанта ( $Magn_i$ ).
14.  $Ver(x, y)$ ,  $y$  — «правильный», «соответствующий назначению», «какой следует» применительно к  $x$ .

15. *Bon* ( $x, y$ ),  $y$  — «хороший» применительно к  $x$ .
16. *Adv<sub>ix</sub>* ( $z, y$ ),  $i = 1, \dots, 4$ ;  $x = A, B, C, D$ ,  $y$  — имя ситуации в роли определения при глаголе, называющем другую ситуацию:  
*Adv<sub>iA</sub>* ( $z, y$ ),  $i = 1, \dots, 4$ ,  $y$  — слово, образованное от  $z$ , которое, заменяя  $z$  в тексте, требует превратить в вершину (вместо  $z$ ) первый актанта этого  $z$  ( $x =$  «сопровождать»,  $y =$  «вместе с»);  
*Adv<sub>iB</sub>* ( $z, y$ ),  $i = 1, \dots, 4$ ,  $y$  требует становиться вершиной второй актанта  $z$  и т.д. ( $x =$  «ошибаться»,  $y =$  «ошибочно»).
17. *Loc* ( $x, y$ ),  $y$  — предлог типовой локализации (пространственной, временной или абстрактной):  
*Loc<sub>in</sub>* ( $x, y$ ),  $y$  — «статическая» локализация (если  $x =$  «Москва», то  $y =$  «в Москве»);  
*Loc<sub>ad</sub>* ( $x, y$ ),  $y$  — предлог направления (если  $x =$  «Москва», то  $y =$  «в Москву»);  
*Loc<sub>ab</sub>* ( $x, y$ ),  $y$  — предлог удаления (если  $x =$  «Москва», то  $y =$  «из Москвы»);  
 Иначе говоря,  $\forall x \forall y (Loc(x, y) \leftrightarrow Loc_{in}(x, y) \vee Loc_{ad}(x, y) \vee Loc_{ab}(x, y))$ .  
 Иногда *Loc* ( $x, y$ ) не удается определить однозначно ( $x =$  «снег»,  $y =$  «на» и  $y =$  «в»).
18. *Copul* ( $x, y$ ),  $y$  — глагол-связка «быть», «являться» ( $x =$  «он напал»,  $y =$  «он совершал нападения»).
19. *Oper<sub>1</sub>* ( $x, y$ ), *Oper<sub>2</sub>* ( $x, y$ ),  $y$  — глагол, связывающий название первого (соответственно второго) актанта в роли подлежащего названием ситуации в роли первого дополнения (если  $x =$  «поддержка», то  $y =$  «оказывать» для *Oper<sub>1</sub>* ( $x, y$ ), а  $y =$  «находить» или «встречать» для *Oper<sub>2</sub>* ( $x, y$ )).
20. *Func<sub>0</sub>* ( $x, y$ ), *Func<sub>1</sub>* ( $x, y$ ), *Func<sub>2</sub>* ( $x, y$ ),  $y$  — глагол, описывающий ситуацию;  $x$  — названия актанта, если они есть, являющихся подлежащим или дополнением ( $x =$  «дождь»,  $y =$  «идти»).
21. *Labor<sub>12</sub>* ( $x, y$ ),  $y$  — глагол, связывающий название первого актанта в роли подлежащего, с названием второго актанта в роли первого дополнения и с названием ситуации в роли второго дополнения ( $x =$  «орден»,  $y =$  «наградать»;  $x =$  «наказание»,  $y =$  «подвергать»).

22.  $Caus_{ij}(x, y)$ ,  $y$  — «делать так, чтобы...», «каузировать» (действие актантов). В случае  $Caus(x, y)$  (без актантных индексов)  $x$  — название неучастника ситуации ( $x$  = «преступление»,  $y$  = «толкать»). Отдельно выступает только при глаголах, в остальных случаях входит в состав сложных параметров.
23.  $Perm_{ij}(x, y) = ne\ Caus_{ij}(x, y)\ ne$  ( $x$  = «преступление»,  $y$  = «не препятствовать»).
24.  $Liqu_{ij}(x, y) = neCaus_{ij}(x, y)$  ( $x$  = «преступление»,  $y$  = «мешать»).
25.  $Incep(x, y)$ ,  $y$  — «начинать». Свойства те же, что и у  $Caus_{ij}(x, y)$ .
26.  $Cont(x, y) = ne\ Incep(x, y)\ ne$ , т.е.  $y$  — «продолжать» = «не переставать».
27.  $Fin(x, y) = Incep(x, y)\ ne$ , т.е.  $y$  — «переставать» = «начинать не».
28.  $Perf(x, y)$  («перфектив»),  $y$  несет завершенность действия, достижение им своего естественного предела. Отдельного самостоятельного выражения  $Perf(x, y)$  в русском языке не имеет; как правило, этот предикат выдает истинное значение, если  $y$  имеет форму совершенного вида.
29.  $Result(x, y)$  («результатив»), т.е.  $y$  — «состояние в результате...»; используется для форм несовершенного вида ( $x$  = «ложиться»,  $y$  = «лечь» для  $Perf(x, y)$ ,  $y$  = «лежать» для  $Result(x, y)$ ).
30.  $Fact^j(x, y)$ ,  $y$  — «реализоваться», «выполниться». Верхний индекс (римские цифры) представляет, если это надо, степень осуществления подразумеваемого требования, причем меньший индекс присваивается более низкой степени (если  $x$  = «капкан» и  $j = I$ , то  $y$  = «срабатывать»; если  $j = II$ , то  $y$  = «поймать»).
31.  $Real^j_{1,2}(x, y)$ ,  $y$  — «реализовать», «выполнить» «требование», содержащееся в  $x$ . Индекс  $j$  имеет то же значение, что и выше — степень выполнения; нижний индекс обозначает глубинно-синтаксический актант, выполняющий требование ( $x$  = «долг (денежный)»,  $y$  = «признать» для  $Real^j_{1,2}(x, y)$ ,  $y$  = «погашать» для  $Real^j_{1,2}(x, y)$ ).
32.  $Prepar^j(x, y)$ ,  $y$  — «приводить в полную готовность к» употреблению, функционированию и т. п. Индекс  $j$ , как и в предикатах  $Fact^j(x, y)$ ,

- $Real_{1,2}^I(x, y)$ , показывает степень готовности ( $x$  = «ружье»,  $y$  = «заряжать») для  $Prepar^I(x, y)$ ,  $y$  = «взводить курок» для  $Prepar^J(x, y)$ .
33.  $Degrad(x, y)$ ,  $y$  — «портиться», «выходить из строя» (=  $IncepPr edAntiVer(x, y)$ : «начинать быть не таким, каким следует» или  $IncepPejor(x, y)$ : «начинать быть хуже»).
34.  $Imper(x, y)$  («императив»),  $y$  — значение «повеления»; для глаголов  $y$  — повелительное наклонение ( $x$  = «уходить»,  $y$  = «брысь» или  $y$  = «кышь» (кошке);  $x$  = «вставать П»,  $y$  = «подъем!» (в армии)).
35.  $Son(x, y)$ ,  $y$  — название типового звучания  $x$  ( $x$  = «корова»,  $y$  = «мычать» или  $y$  = «му-у-у»).
36.  $Destr(x, y)$ ,  $y$  — типовое название «агрессивного» действия ( $x$  = «оса»,  $y$  = «жалит»).
37.  $Cap(x, y)$ ,  $y$  — «начальник» ( $x$  = «факультет»,  $y$  = «декан»).
38.  $Equip(x, y)$ ,  $y$  — «личный состав» ( $x$  = «население»,  $y$  = «государства»).
39.  $Doc(x, y)$ ,  $y$  — «документ»:  
 $Doc_{res}(x, y)$ ,  $y$  — «документ», являющийся результатом («воплощающий в себе»;  $x$  = «отчитываться»,  $y$  = «отчет»);  
 $Doc_{perm}(x, y)$ ,  $y$  — «документ на право...» ( $x$  = «поезд»,  $y$  = «(проездной) билет» для  $Doc_{perm}Oper_2(x, y)$ );  
 $Doc_{cert}(x, y)$ ,  $y$  — «документ, удостоверяющий...» ( $x$  = «высшее образование»,  $y$  = «диплом»);  
 Иначе говоря,
40.  $Attr(x, y)$ ,  $y$  — типовая метонимия для  $x$  ( $x$  = «школьник»,  $y$  = «парта»;  $x$  = «офицер»,  $y$  = «погоны»).
- Помимо перечисленных выше простых лексических предикатов для описания лексической сочетаемости могут использоваться и их комбинации — составные предикаты:
- $AntiReal_2(x, y)$ : проваливать экзамен /проваливаться на экзамене.  
 $IncepOper_1(x, y)$ : приобретать популярность, впадать в отчаяние.  
 $IncepOper_2(x, y)$ : поступать в продажу, попадать под обстрел.  
 $CausOper_2(x, y)$ : ставить под контроль, пускать в обращение.

В общем случае лексическая функция определяется не для всех слов и словосочетаний.

Во-первых, некоторые лексические функции определены лишь для той или иной части речи; так, *Oper*, *Func* и *Labor* мыслимы лишь для существительных, а *Incep*, *Cont* и *Fin* — лишь для глаголов. В случае, когда лексические функции представляются в виде предикатов, то это не вызывает затруднений. В данном случае соответствующие предикаты ложны.

Во-вторых, та или иная функция может определяться только для слов определенной семантики. Например: *Magn* — для слов, смысл которых допускает градацию «больше — меньше»; *Cap* и *Equip* — для слов, смысл которых предполагает наличие «начальника» и «персонала», т. е. для названий учреждений и организаций в самом широком понимании. Далее, *Conv* — для слов, называющих отношения с двумя и более местами; *Real* — для слов, в смысл которых входит компонент «требовать» («нужно») и т. д. Как уже отмечалось, *Oper*, *Func* и *Labor* определены только для названий ситуаций.

Следует иметь в виду, что и при вполне подходящем (по своим синтаксическим и семантическим свойствам) аргументе лексическая функция может не иметь значения (в данном языке). Например, синонимы, в принципе, возможны для любых слов, а имеются только у некоторых: при *Oper<sub>1</sub>* (*визит*) = офиц. *наносить* [~ Сдат], *являться* (*приходить*) [*с* ~ом к Сдат], *Oper<sub>2</sub>* для слова *визит* в современном русском языке отсутствует (по смыслу должно было бы быть что-то вроде \**получать*, \**принимать*, \**подвергаться*).

## 2. ВАЛЕНТНОСТИ

Валентностями обладают слова, которые являются предикатами, т. е. те, которые задают ситуацию. Это все глаголы, некоторые существительные (отглагольные), прилагательные (обозначающие сравнение: больше, меньше, выше, ниже), некоторые предлоги и наречия.

Валентности слова бывают синтаксические и семантические. Семантические валентности определяются лексическим анализом ситуации, задаваемой этим словом. Приведём пример со словом *аренда* или *арендовать*. *А арендует* С значит, в первом приближении, что за какое-то вознаграждение D лицо А приобретает у другого лица В право на эксплуатацию недви-

жимой собственности  $C$  в течении времени  $T$ . Следовательно, существенными для ситуации аренды являются следующие «участники» или семантические актанты: субъект аренды (тот, кто арендует), первый объект аренды (то, что арендуют), контрагент (тот, у кого арендуют), второй объект (плата) и срок.

Эти актанты необходимы, так как устранение какого-либо из них изменяет смысл ситуации. Например если убрать срок, то ситуация аренды трансформируется в ситуацию купли-продажи. С другой стороны, эти актанты достаточны, поскольку в ситуации аренды не требуется указание того, по какой причине, где, когда и с какой целью она осуществлялась. Хотя соответствующие словоформы грамматически присоединимы к глаголу *арендовать*.

В итоге, эта ситуация имеет 5 валентностей и формально записывается в виде предиката  $P(x_1, x_2, x_3, x_4, x_5)$ , где  $x_1$  — «кто»,  $x_2$  — «что»,  $x_3$  — «у кого»,  $x_4$  — «цена»,  $x_5$  — «срок».

В предложении могут быть определены актанты не для всех семантических валентностей, некоторые могут просто не упоминаться или вообще не иметь синтаксического выражения.

Синтаксические валентности — это те, которые представлены в тексте. Они определяются присоединяемыми к слову подлежащими и дополнениями и зависят от контекста.

Например, глагол *промахнуться* имеет 4 семантические валентности: кто (деятель), во что/по чему (мишень), из чего (оружие — факультативно) и чем (орган, средство). Но в большинстве контекстов синтаксически выражается лишь первая валентность. Например, нельзя сказать «*Он промахнулся в окно бутылкой*».

Возможны случаи, когда синтаксических валентностей у слова больше, чем семантических.

Для того чтобы не связывать с каждым глаголом (и другими словами) отдельный предикат, будем рассматривать предикат, размерность которого больше на 1:  $P^{val}(y, x_1, x_2, \dots, x_n)$ , при этом  $y$  будет само слово, а  $x_1, x_2, \dots, x_n$  — его валентности. Чтобы отличать синтаксические и семантические актанты, можно использовать мультииндексы с целью указать, какие актанты заданы в тексте. Запись  $P_{i_1 i_2 \dots i_k}^{val}(y, x_{i_1}, x_{i_2}, \dots, x_{i_k})$  означает, что заданы актанты  $i_1, i_2, \dots, i_k$ . В частности если заданы все актанты, то получаем  $P_{1 \dots n}^{val}(y, x_1, x_2, \dots, x_n)$ . Некоторые варианты (наборов мультииндексов) могут

быть недопустимы в языке. Если набор  $i_1, i_2, \dots, i_k$  допустим, то имеет место импликация:

$$\forall y \forall x_1 \dots \forall x_n (P_{1..n}^{val}(y, x_1, x_2, \dots, x_n) \rightarrow P_{i_1 i_2 \dots i_k}^{val}(y, x_{i_1}, x_{i_2}, \dots, x_{i_k})).$$

Более того, если имеется два набора допустимых мультииндексов  $\langle i_1, i_2, \dots, i_k \rangle$  и  $\langle i'_1, i'_2, \dots, i'_s \rangle$  таких, что  $\{i_1, i_2, \dots, i_k\} \supseteq \{i'_1, i'_2, \dots, i'_s\}$ , то имеет место аналогичная импликация:

$$\forall y \forall x_{i_1} \dots \forall x_{i_k} \forall x_{i'_1} \dots \forall x_{i'_s} (P_{i_1 \dots i_k}^{val}(y, x_{i_1}, x_{i_2}, \dots, x_{i_k}) \rightarrow P_{i_1 i_2 \dots i_s}^{val}(y, x_{i_1}, x_{i_2}, \dots, x_{i_s})).$$

### 3. ТОЛКОВО-КОМБИНАТОРНЫЙ СЛОВАРЬ

Можно считать, что язык — это очень большая модель, в которой определены лексические предикаты, действующие соответствующим, описанным выше образом.

Статья толково-комбинаторного словаря несет информацию о валентностях конкретного слова, верную не только в ее рамках, но и в рамках всего языка в целом. Валентности соответствует предикат  $P^{val}(c_x, \bar{y})$ , где  $\bar{y} = y_1, \dots, y_n$  — семантические актаны слова  $c_x$ ,  $n$  — валентность слова  $c_x$ . Например, в предложении *Петя читает книгу* будет  $c_x = \text{"читать"}$ ,  $n = 2$ :  $y_1 = \text{"Петя"}$ ,  $y_2 = \text{"книга"}$ , т.е. условно можно написать  $P^{val}(c_x, y_1, y_2) = 1$ .

Набор статей в толково-комбинаторном словаре можно считать некоторой подмоделью исходной модели, являющейся языком. Лексические предикаты, определенные теперь на более узком множестве, будут действовать аналогично.

Пусть  $\Phi$  — множество правильно построенных фраз языка  $L$ , и  $\varphi \in \Phi$  — фраза из этого множества,  $c_x \prec \varphi$  — слово  $c_x$  входит во фразу  $\varphi$ , причем  $c_x \in L$ , а  $c_x$  — существительное или прилагательное. Обозначим Predicate — множество предикатов, определенных на  $L$ . Одним из элементов этого множества является введенный ранее предикат валентности  $P^{val}(c_x, y_1, \dots, y_n)$ .

Аналогично, можно предполагать, что имеются другие предикаты:

- ♦ предикат рода слова  $Case(c_x, \overline{Cas})$ , где  $\overline{Cas} \in \{cas_1 = "жен."; cas_2 = "муж."; cas_3 = "ср." \}$ ;
- ♦ предикат предлога  $Preposition(c_x, \overline{Prep})$ , где  $\overline{Prep} \in \{pr_1, \dots, pr_k \}$  — множество предлогов, сочетаемых с данным словом;
- ♦ предикат падежа  $Gender(c_x, \overline{G})$ , где  $\overline{G}$  — падеж слова  $c_x$ . Для разных языков число падежей различно. Например, для русского языка имеется шесть падежей:  
 $g_1 = "им.н."; g_2 = "род.н."; g_3 = "дат.н."; g_4 = "вин.н.";$   
 $g_5 = "твор.н."; g_6 = "предл.н."$ ,  
 для немецкого — четыре падежа:  
 $g_1 = "N"; g_2 = "G"; g_3 = "D"; g_4 = "A"$ .

Словарная статья толково-комбинаторного словаря имеет следующий вид. Она содержит основное слово, лексические предикаты, связанные с ним, и информацию о валентности данного слова.

Информация о валентности включает в себя число, указывающее число актантов, и для каждого актанта — указание, в каких падежах и с какими предлогами используются слова, соответствующие данному актанту. В отдельных случаях может быть указан также род слова.

Сказанное выше может быть представлено посредством набора предикатов вида

$$P(x_i, \overline{Cas}, \overline{Prep}, \overline{G}) = Case(x_i, \overline{Cas}) \& Preposition(x_i, \overline{Prep}) \& Gender(x_i, \overline{G}),$$

где  $x_i$  — свободная переменная, соответствующая  $i$ -му актанту.

#### 4. ЧАСТИ РЕЧИ И СИНТАКСИЧЕСКИЕ ТИПЫ, СТРУКТУРНЫЕ МЕТОДЫ МАРКУСА

##### 4.1. Основные определения и обозначения

Назовем конечное множество слов *словарем*  $\Gamma$ . Рассмотрим свободную полугруппу  $T$  на  $\Gamma$ , а именно, множество всех конечных последовательностей слов с определенной на нем ассоциативной и некоммутативной бинарной операцией конкатенации (сцепления). Последовательность слов мы бу-

дем называть также последовательностью (цепочкой) над  $\Gamma$ . Нулевая последовательность, обозначим ее  $\theta$ , — это такая последовательность, что  $\theta x = x\theta = x$  для каждой последовательности  $x$ . Если это специально не оговорено,  $\theta$  не принадлежит  $\Gamma$ .

Подмножество  $\Phi \subseteq T$  назовем *языком* над  $\Gamma$ . Полугруппа  $T$  в целом будет называться в этом случае *полным* или *универсальным языком* над  $\Gamma$ .

Тройку  $\{\Gamma, P, \Phi\}$ , где  $\Gamma$  — конечный словарь,  $P$  — разбиение  $\Gamma$ , а  $\Phi$  — подмножество свободной полугруппы над  $\Gamma$ , мы назовем *языком с парадигматической структурой* или просто *языком*.

При машинном переводе требуется рассматривать языки с более сложной структурой. При этом в качестве языка должна рассматриваться система  $\{\Gamma, P, \Phi, K, \varphi\}$ , где  $\Gamma$ ,  $P$  и  $\Phi$  — определенные выше объекты,  $K$  — класс подмножеств  $\Gamma$ , называемых грамматическими категориями (например, множество слов в именительном падеже и множество слов в прошедшем времени), а  $\varphi$  — функция, сопоставляющая каждому слову  $x$  пересечение всех грамматических категорий, содержащих  $x$ .

Рассмотрим наиболее простое понятие языка, который в этом случае задается как пара  $\{\Gamma, \Phi\}$ . Последовательности, принадлежащие  $\Phi$ , мы назовем *отмеченными последовательностями*.

Наиболее важным разбиением  $\Gamma$ , которое встречается в лингвистике, является так называемое разбиение на *дистрибутивные классы*. Оно определяется следующим образом. Два слова  $a$  и  $b$  принадлежат одному дистрибутивному классу, если для каждой пары последовательностей  $x, y$  из  $\Phi$  следует  $xby \in \Phi$ , и обратно.

Введем понятие *контекста*. Контекст может быть определен как упорядоченная пара последовательностей над  $\Gamma$ . Мы будем обозначать контекст  $\langle x, y \rangle$ , где  $x \in T$  и  $y \in T$ . Слово  $a$  *допустимо* в контексте  $\langle x, y \rangle$ , если последовательность  $xay$  принадлежит  $\Phi$ . Обозначим через  $\mathbf{J}(a)$  множество всех контекстов, в которых допустимо  $a$ . Отсюда непосредственно следует, что два слова  $a$  и  $b$  принадлежат одному дистрибутивному классу тогда и только тогда, когда  $\mathbf{J}(a) = \mathbf{J}(b)$ , т.е. тогда и только тогда, когда  $a$  и  $b$  допустимы в одних и тех же контекстах.

Легко видеть, что два различных дистрибутивных класса не пересекаются; таким образом, эти классы определяют разбиение  $\Gamma$ , которое мы назовем *дистрибутивным разбиением*  $\Gamma$ . Дистрибутивный класс можно назы-

вать также *семейством*. Семейство, содержащее  $a$ , мы будем обозначать  $S(a)$ . Мы будем употреблять эти два термина как эквивалентные.

Более сложным является понятие языка как тройки  $\{\Gamma, P, \Phi\}$ , где  $P$  — разбиение  $\Gamma$ , отличное от разбиения на дистрибутивные классы. Если  $P$  — разбиение  $\Gamma$ , то каждое подмножество  $P$  мы будем называть *клеткой из  $P$*  или  *$P$ -клеткой*. Если определено разбиение  $P$ , то мы можем написать:

$\Gamma = \bigcup_{i=1}^n P_i$ , где  $P_i$  обозначает клетку из  $P$ , а число клеток равно  $n$ . Поскольку

подмножества  $P_i$  не пересекаются, каждое слово принадлежит одной и только одной клетке. Обозначим через  $P(a)$  клетку из  $P$ , содержащую слово  $a$ . Из сказанного следует, что для двух различных слов  $a$  и  $b$  выполняется либо  $P(a) = P(b)$ , либо  $P(a) \cap P(b) = 0$ .

Множество  $P(a)$  обычно интерпретируется в естественных языках как множество флективных форм слова  $a$ . Такая интерпретация требует введения так называемого *единичного разбиения*  $\Gamma$ , в котором каждая клетка состоит из одного слова.

Рассмотрим два разбиения  $P$  и  $Q$  словаря  $\Gamma$ . Будем говорить, что разбиение  $P$  *мельче* разбиения  $Q$ , если  $P(a) \subseteq Q(a)$  для каждого  $a \in \Gamma$ .

Единичное разбиение является самым мелким разбиением  $\Gamma$ , а любое разбиение  $\Gamma$  мельче несобственного разбиения. Если мы интерпретируем  $P(a)$  как множество всех флективных форм  $a$ , то разбиение  $P$  окажется мельче, чем разбиение  $\Gamma$  на части речи.

Если  $x_1 x_2 \dots x_n$  — последовательность над  $\Gamma$ , то последовательность  $P(x_1)P(x_2)\dots P(x_n)$  называется  *$P$ -структурой* последовательности  $x_1 x_2 \dots x_n$ .  *$P$ -структура* является *отмеченной*, если последовательность  $x_1 x_2 \dots x_n$  может быть выбрана так, что она принадлежит  $\Phi$ . Другими словами,  *$P$ -структура*  $P_1 P_2 \dots P_s$  отмечена, если существует отмеченная последовательность  $x_1 x_2 \dots x_n$  такая, что  $P_i = P(x_i)$  при  $1 \leq i \leq s$ .

Далее будем говорить, что  $P$  *регулярно мельче*  $Q$ , если  $P$  мельче  $Q$ , и для каждой тройки слов  $x, y, z$  выполнение включений  $P(x) \subseteq Q(z)$  и  $P(y) \subseteq Q(z)$  влечет  $P$ -эквивалентность  $P(x) \leftrightarrow P(y)$ .

В качестве простейшего примера регулярно более мелкого разбиения возьмем единичное разбиение  $E$ . Это разбиение регулярно мельче, чем разбиение  $S$  на дистрибутивные классы. Для каждого разбиения  $P$  над  $\Gamma$  рас-

смотрим разбиение  $P'$ , клетки которого определены следующим образом:  $P'(x) = \bigcup_{P(x) \leftrightarrow P(y)} P(y)$  (для каждого  $x \in \Gamma$ ), где объединение берется по

всем словам  $y$ , для которых  $P(y) \leftrightarrow P(x)$ .

По своему определению разбиение  $P'$  таково, что  $P$  регулярно мельче  $P'$ . Разбиение  $P'$  назовем *производным* от разбиения  $P$ . Следует отметить, что разбиение  $S$  на дистрибутивные классы является производным от единичного разбиения  $E$ :  $S = E'$ . В языке  $\{\Gamma, P, \Phi\}$  производное  $P'$  определено однозначно. Но может существовать несколько языков вида  $\{\Gamma, P, \Phi\}$ , для которых производное будет одинаковым.

Рассмотрим язык  $\{\Gamma, P, \Phi\}$ . *Цепью между словами  $a$  и  $b$*  назовем конечную последовательность  $x_1, x_2, \dots, x_i, x_{i+1}, \dots, x_n$  такую, что  $x_1 = a$ ,  $x_n = b$  и  $x_i \in S(x_{i+1}) \cup P(x_{i+1})$  для  $1 \leq i \leq n-1$ .

Число  $n$  определяет длину цепи. Будем считать, что для каждого слова  $a$  существует цепь длиной 1, соединяющая  $a$  с самой собой.

Обозначим  $R(a)$  множество таких слов  $b$ , для каждого из которых существует цепь, соединяющая  $a$  и  $b$ . Как легко видеть, (1)  $a \in R(a)$ ; (2) если  $b \in R(a)$ , то  $a \in R(b)$ ; (3) если  $b \in R(a)$  и  $c \in R(b)$ , то  $c \in R(a)$ . Таким образом, множества  $R(a)$  при  $a \in \Gamma$  определяют разбиение  $R$  словаря  $\Gamma$ . Такое разбиение мы назовем *разбиением из смешанных клеток*. Для каждого  $a \in \Gamma$  выполняется  $S(a) \subseteq R(a)$  и  $P(a) \subseteq R(a)$ . Отсюда видно, что разбиения  $P$  и  $S$  мельче  $R$ . Объединение  $P$  и  $S$  составляет  $R$ .

Обозначим через  $K(a)$  множество слов  $b$  таких, что выполняется по крайней мере одно из двух следующих условий:  $P(a) \cap S(b) \neq \emptyset$  (1);  $P(b) \cap S(a) \neq \emptyset$  (2). Множество  $K(a)$  по определению является классом  $a$ . Поскольку  $a \in P(a) \cap S(a)$ , то следовательно,  $a \in K(a)$  для каждого  $a \in \Gamma$ .

## 4.2. Типология языков

Язык  $\{\Gamma, P, \Phi\}$  называется *адекватным*, если для каждого  $x \in \Gamma$  выполняется включение  $S(x) \subseteq P'(x)$ . Простейший пример адекватного языка — язык, в котором  $P$  представляет собой единичное разбиение  $\Gamma$ .

Для дальнейших рассуждений нам пригодятся следующие теоремы.

**Теорема 1.** Существует неадекватный язык.

**Теорема 2.** Если  $\{\Gamma, P, \Phi\}$  — адекватный язык, то  $R' = P'$ .

**Теорема 3.** Пусть  $\{\Gamma, P, \Phi\}$  — адекватный язык. Если классы  $K(x)$  определяют разбиение  $K$  словаря  $\Gamma$ , то  $K' = P'$ .

**Теорема 4.** Если язык  $\{\Gamma, P, \Phi\}$  однороден, то  $K(x) = R(x)$  для каждого  $x \in \Gamma$ , т.е. классы совпадают со смешанными клетками.

### 4.3. Части речи и синтаксические типы

Рассмотрим язык  $\{\Gamma, P, \Phi\}$ , интерпретируя  $\Gamma$  как словарь естественного языка  $L$ ,  $P(x)$  (при  $x \in \Gamma$ ) — как множество всех форм слова  $x$ , а  $\Phi$  — как множество всех правильно построенных предложений языка  $L$ . Если слово  $x$  имеет две различные формы  $x_1$  и  $x_2$ , то будем считать, что  $P(x_1) \cap P(x_2) = \emptyset$ . Две омонимичные формы, такие, как *free* (прилагательное) и *free* (глагол), будут рассматриваться как различные слова, а соответствующие им  $P$ -клетки — как непересекающиеся.

При предлагаемой интерпретации клетки производного разбиения  $P'$  могут рассматриваться как приближенная модель частей речи в языке  $L$ .

Легко заметить, что для каждого  $a \in \Gamma$   $P$ -клетка  $P(a)$  содержит только формы  $a$ . Для некоторых слов  $a$   $P(a)$  не содержит всех форм  $a$ , а только те, которые встречаются в предположениях из  $\Phi$ , но в любом естественном языке выполняется одно очень простое условие: если два слова  $a$  и  $b$  принадлежат одной парадигме, т.е. если  $b \in P(a)$ , то они принадлежат одной части речи. В соответствии с этим правилом любое строгое определение понятия «часть речи» должно рассматривать любую часть речи как объединение  $P$ -клеток.

Пусть даны два слова  $a$  и  $b$ ; мы можем считать, что  $P(a)$  и  $P(b)$  относятся к одной части речи тогда и только тогда, когда  $P(a)$  и  $P(b)$   $P$ -эквивалентны, т.е. тогда и только тогда, когда  $b \in P'(a)$ .

Части речи не имеют абсолютного характера, они зависят от множества  $\Phi$  отмеченных последовательностей (т.е. от синтаксиса языка), от разбиения  $P$  (т.е. от морфологии языка). Части речи в естественном языке являются функцией  $P$  и  $\Phi$ , но мы никогда не учитываем при этом всех возможных предложений и парадигм данного языка. В каждом конкретном случае мы рассматриваем его фрагмент, достаточно сложный, чтобы получить хо-

рошее приближение к естественному языку, но в то же время достаточно простой, чтобы было возможно его систематическое и детальное исследование.

Объясним причины, побудившие дать частям речи приведенное выше определение. Рассмотрим язык  $L = \{\Gamma, P, \Phi\}$ . Этому языку мы можем поставить в соответствие другой язык, а именно  $P(L) = \{\Gamma_1, P_1, \Phi_1\}$ , где  $\Gamma_1$  — множество всех  $P$ -клеток языка  $L$ ,  $P_1$  — разбиение  $\Gamma_1$  на  $P$ -эквивалентные классы, а  $\Phi_1$  — множество всех отмеченных  $P$ -структур в языке  $L$ . Язык  $P(L)$  будет называться  $P$ -абстракцией языка  $L$ . Данный уровень абстракции является именно таким, на котором становится понятной логическая структура частей речи, поскольку в  $P(L)$  они не что иное, как дистрибутивные классы. Таким образом,  $P'$ -клетки языка  $L$  в точности соответствуют дистрибутивным классам языка  $P(L)$ .

Пусть дана часть речи  $\mathbf{P}$ . Интересно найти класс  $\mathbf{C}$  контекстов, для которых выполняются следующие условия: 1) для каждого слова  $a \in \mathbf{P}$  существует контекст  $(x, y) \in \mathbf{C}$  такой, что  $xa y \in \Phi$ ; 2) если дано слово  $b$ , которое не принадлежит  $\mathbf{P}$ , то для такого слова не существует контекста  $(u, v) \in \mathbf{C}$  такого, что  $ubv \in \Phi$ . Класс  $\mathbf{C}$  называется *диагностическим классом* для  $\mathbf{P}$ .

Выделение диагностических классов является очень важной задачей структурной лингвистики, поскольку диагностические классы для  $\mathbf{P}$  позволяют нам выделить и изучать различные  $\mathbf{P}$  по отдельности.

Выше не делалось никаких предположений о природе рассматриваемого языка. Но существует общая гипотеза, которая утверждает, что каждый естественный язык является адекватным.

Вспомним теорему 2 из предыдущего раздела. Если  $\{\Gamma, P, \Phi\}$  — адекватный язык, то  $R' = P'$ . Эта теорема дает возможность по-новому определить понятие частей речи. Для данного слова  $a \in \Gamma$  его часть речи совпадает с  $R'(a)$ . Еще одна возможность определения понятия «часть речи» в случае адекватного языка дается теоремой 3. Если в таком языке классы  $K(x)$  определяют разбиение  $\Gamma$ , то  $K' = P'$ . Отсюда часть речи, которой принадлежит слово  $a$ , совпадает с  $K'(a)$ . Таким образом, становится ясно, что существуют адекватные неоднородные языки, в которых части речи могут быть построены тремя способами: как множество  $P'$ -клеток,  $R'$ -клеток и  $K'$ -клеток. Этот факт очень важен, поскольку в естественных языках суще-

ствуется много фрагментов, которые являются адекватными, но не однородными.

Наиболее благоприятные условия в отношении анализа понятия «часть речи» предоставляют однородные языки. Действительно, согласно теореме 4 в любом однородном языке классы совпадают со смешанными клетками, т.е.  $K(x) = R(x)$  для любого  $x \in \Gamma$ . Поскольку в любом адекватном языке для каждого  $x \in \Gamma$  выполняется  $R'(x) = P'(x)$ , то, следовательно, в однородном языке  $K'(x) = R'(x) = P'(x)$  для любого  $x \in \Gamma$ .

В соответствии с теоремой 4 можно построить части речи, взяв производное разбиение  $K'$ . Если  $a$  и  $b$  — два существительных различных родов, а  $c$  — прилагательное, то  $K$ -структуры  $K(a)$  и  $K(b)$   $K$ -эквивалентны. Таким образом,  $K(a)$  содержит все существительные. Если  $a$  — существительное, а  $c$  — прилагательное, то  $K(a)$  и  $K(c)$  не являются  $K$ -эквивалентными. Значит, части речи в том виде, как они описываются приведенной схемой, совпадают с традиционными.

Различия могут возникнуть, если иначе выбрать  $\Gamma$  и  $\Phi$ ; в особенности это касается местоимений, артиклей, числительных, наречий и некоторых типов прилагательных. Так, например, если не считать правильно построенным французским предложением такую последовательность слов, как *il est très mort*, то слово *mort*, не может быть отнесено к той же самой  $P'$ -клетке, что и *beau*.

Пусть даны два языка  $L_1 = \{\Gamma_1, P_1, \Phi_1\}$  и  $L_2 = \{\Gamma_2, P_2, \Phi_2\}$ . Будем говорить, что они являются  $P'$ -изоморфными, если существует взаимно однозначное отображение  $\varphi$  словаря  $\Gamma_1$  на  $\Gamma_2$  такое, что  $y \in P'_1(x)$  в  $L_1$  тогда и только тогда, когда  $\varphi(y) \in P'_2(\varphi(x))$  в  $L_2$ . Поскольку части речи в точности совпадают с  $P'$ -клетками данного языка, то следовательно,  $P'$ -изоморфизм сохраняет подразделение на части речи.

Сущность проблемы синтаксических типов была сформулирована Ламбеком с помощью следующей аналогии: в классической физике можно просто проверить «грамматическую правильность» уравнения, сравнивая размерности его правой и левой частей. Спрашивается, нельзя ли подобным же образом приписывать грамматические типы словам в естественном языке так, чтобы грамматическая правильность предложения могла бы быть определена с помощью операций над этими типами.

Рассмотрим словарь  $V$ . Начнем с приписывания нескольких простейших типов отдельным словам и последовательностям над  $V$ . Из этих основных

типов с помощью трех формальных операций можно построить сложные типы. Такими операциями будут: *умножение*, *левое деление* и *правое деление*; первая из этих операций не имеет специального значка, а левое и правое деления обозначаются соответственно  $\backslash$  и  $/$ . Выражение  $X \rightarrow x$  означает, что последовательность  $X$  относится к типу  $x$ . Сложные типы определяются следующим образом: если  $X \rightarrow x$  и  $Y \rightarrow y$ , то  $XY \rightarrow xy$ ; если  $XY \rightarrow z$  и  $Y \rightarrow y$ , то  $X \rightarrow z/y$  (читается  $z$  над  $y$ ); если  $XY \rightarrow z$  и  $X \rightarrow x$ , то  $Y \rightarrow x \backslash z$  (читается  $x$  под  $z$ ). Другими словами, выражение типа  $x/y$ , за которым следует выражение типа  $y$ , приводит к выражению типа  $x$ , так же как и выражение типа  $y \backslash x$ , которому предшествует выражение типа  $y$ .

Если любое выражение типа  $x$  принадлежит также типу  $y$ , то пишут  $x \rightarrow y$ . Из определения левого и правого делений следует, что и  $(x/y)y \rightarrow x$ , и  $y(y \backslash x) \rightarrow x$ .

В некоторых случаях, когда мы имеем дело с фрагментами естественных языков, мы можем рассматривать только два элементарных типа:  $s$  — тип предложений (т.е. отмеченных последовательностей) и  $n$  — именной тип. Но в более сложных случаях приходится учитывать значительно большее число элементарных типов. Так, иногда мы вводим в качестве элементарного типа  $i$  — инфинитив непереходного глагола. Для простоты ограничим здесь использование типа  $s$  и будем приписывать его только полным изъявительным предложениям (таким образом, мы исключаем большинство элементов диалога).

В наших примерах под именем будем понимать, прежде всего, собственные имена, а также такие выражения, которые могут встретиться в контекстах, аналогичных тем, в которых выступают собственные имена. Так, тип  $n$  приписывается так называемым неисчисляемым именам: *milk* «молоко», *rice* «рис» и т.д., которые могут встретиться без артикля, а также именным словосочетаниям типа *poor John* «бедный Джон» или *fresh milk* «свежее молоко». Но мы не можем приписать такой тип так называемым исчисляемым именам таким, как *king* «король», *chair* «стул» и т.д., которые требуют перед собой артикль, а также местоимению *he*, поскольку последнее не может заменять слова *John*, например, в предложении *poor John works*.

Рассмотрим словарь  $\Gamma$  и множество элементов, которые будем называть *элементарными типами*. Предположим теперь, что взята некоторая последовательность над  $\Gamma$ , и ее элементам приписаны определенные элементарные типы. Определим теперь множество  $\Gamma$  элементов, называемых *типами*, следующим образом: все элементарные типы являются типами; если  $x$  и  $y$  —

типы, то  $xy$ ,  $x \setminus y$  и  $x/y$  — также типы. Основные правила исчисления имеют следующий вид:

- (1)  $x \rightarrow x$  ;
- (2)  $(xy)z \rightarrow x(yz)$  ;
- (2')  $x(yz) \rightarrow (xy)z$  ;
- (3) если  $xy \rightarrow z$ , то  $x \rightarrow z/y$  ;
- (3') если  $xy \rightarrow z$ , то  $y \rightarrow x \setminus z$  ;
- (4) если  $x \rightarrow z/y$ , то  $xy \rightarrow z$  ;
- (4') если  $y \rightarrow x \setminus z$ , то  $xy \rightarrow z$  ;
- (5) если  $x \rightarrow y$  и  $y \rightarrow z$ , то  $x \rightarrow z$ .

В соответствии с видом правил (2) и (2') предлагаемое синтаксическое исчисление можно назвать *ассоциативным синтаксическим исчислением*. Его можно рассматривать абстрактно как формальный язык или как систему вывода. Естественно, есть дополнительное множество правил, которое имеет место в данном ассоциативном синтаксическом исчислении. Но некоторые правила здесь оказываются неверными. Например, следующие правила неверны:  $(x/y)/z \rightarrow x/(y/z)$ ,  $(x/y) \setminus z \rightarrow x/(y \setminus z)$ ,  $xy \rightarrow yx$ ,  $z \rightarrow (z/y)y$ .

До сих пор типы сопоставлялись любым последовательностям слов, а не только тем, которые имеют завершённую форму. Предположим теперь, что мы приписываем типы не просто последовательностям слов, а грамматически правильным словосочетаниям, т.е. как бы расставляя в последовательности слов (и даже в последовательности морфем) скобки.

Рассмотрим множество последовательностей, называемых *атомарными словосочетаниями*, и примем следующее рекурсивное определение *словосочетания* (последовательное написание букв соответствует операции конкатенации, скобки имеют собственное значение, внешние скобки фразы могут быть опущены): все атомарные словосочетания являются словосочетаниями; если  $A$  и  $B$  — словосочетания, то  $(AB)$  — также словосочетание.

Типы вводятся с помощью аналогичного рекурсивного определения. Рассмотрим конечное множество, элементы которого называются *элементарными типами*. Все элементарные типы являются типами. Определены три бинарные операции над типами так, что если  $x$  и  $y$  — типы, то  $(xy)$ ,  $(x/y)$  и  $(x \setminus y)$  — также типы.

Типы сопоставляются словосочетаниям по следующим правилам: если  $A$  относится к типу  $a$  и  $B$  — к типу  $b$ , то  $(AB)$  — к типу  $(ab)$ ; если  $(AB)$  относится к типу  $c$  и все  $B$  — к типу  $b$ , то  $A$  относится к типу  $(c/b)$ ; если  $(AB)$  относится к типу  $c$ , а все  $A$  — к типу  $a$ , то  $B$  относится к типу  $(a \setminus c)$ . Система основных правил для неассоциативного синтаксического исчисления имеет вид, аналогичный системе правил для ассоциативного синтаксического исчисления, за исключением того, что не выполняются правила  $(xy)z \rightarrow x(yz)$  (2) и  $x(yz) \rightarrow (xy)z$  (2').

Мы можем теперь рассматривать неассоциативное синтаксическое исчисление как систему вывода. Рассмотрим сначала множество элементов, которые называются *переменными*. Затем рекурсивно определим еще одно множество, элементы которого назовем *термами*: все переменные являются термами; если  $x$  и  $y$  — термы, то  $(xy)$ ,  $(x \setminus y)$  и  $(x / y)$  — также термы. Введем единственную *формулу*:  $x \rightarrow y$  (где  $x$  и  $y$  — термы), единственную *аксиому*:  $x \rightarrow x$  и правила вывода (3), (3'), (4), (4') и (5). Синтаксическое исчисление позволяет нам перевести некоторые грамматические правила из грамматики в словарь.

Предположим, что мы заменили все грамматические правила языка, сопоставив подходящим образом типы единицам словаря. При этом становится возможным проанализировать данную последовательность слов автоматически. Расставляя скобки, мы выделяем в последовательности словосочетания и подписываем под каждым словом один из типов, сопоставленных ему в словаре. Пусть  $x$  — сложный тип, соответствующий всей последовательности. Если имеем  $x \rightarrow s$ , то  $X$  — утвердительное предложение, если  $x \rightarrow i$ , то  $X$  — повелительное предложение, и т.д. Этот процесс повторяется для всех вариантов расстановки скобок и всех вариантов приписывания словам типов.

Приведем формальное определение категориальных грамматик. *Двунаправленной категориальной грамматикой* называется кортеж  $\langle \Gamma, C, \Sigma, R, f \rangle$  из следующих пяти элементов:  $\Gamma$  — конечное множество элементов (словарь),  $C$  — замыкание конечного множества основных категорий, обозначим их  $\psi_1, \dots, \psi_n$  относительно операции левого и правого делений (таким образом, если  $\alpha$  и  $\beta$  — категории, то  $\alpha / \beta$  и  $\alpha \setminus \beta$  — также категории),  $\Sigma$  — выделенная в множестве  $C$  категория (категория предложений),  $R$  — множество, состоящее из двух правил сокращения:  $(\varphi_i / \varphi_j)\varphi_j \rightarrow \varphi_i$  и

$\varphi_i(\varphi_i \setminus \varphi_j) \rightarrow \varphi_j, f$  — функция, сопоставляющая элементам из  $\Gamma$  конечные подмножества из  $C$ .

Мы будем говорить, что последовательность категорий  $\beta$  непосредственно сокращается до  $\alpha$ , если  $\beta$  является результатом применения к  $\alpha$  одного из правил сокращения. Мы будем также говорить, что  $\alpha$  сокращается до  $\beta$ , если  $\beta$  является результатом применения конечного числа правил сокращения к  $\alpha$ ; более точно, если существует последовательность категорий  $\gamma_1, \dots, \gamma_n$  такая, что  $\alpha = \gamma_1, \beta = \gamma_n$  и  $\gamma_i$  непосредственно сокращается до  $\gamma_{i+1}$  при  $i = 1, \dots, n-1$ .

Последовательность  $x$  над  $\Gamma$  называется предложением тогда и только тогда, когда по крайней мере одна последовательность категорий, приписанных элементам  $x$  в соответствии с  $f$ , сокращается до  $\Sigma$ . Множество всех таких предложений образует язык, задаваемый данной двунаправленной категориальной грамматикой. Такой язык называется двунаправленным категориальным языком.

#### 4.4. Грамматический род

Грамматический род представляет собой одну из наиболее интересных проблем теории грамматики. Исследовались различные аспекты этой проблемы, такие как связь между значением существительного и его родом; соотношение между родом существительного и его окончанием. Род рассматривался в свете соответствия между планом содержания и планом выражения с точки зрения синтаксических и контекстуальных связей, с точки зрения его происхождения и эволюции.

Почти все авторы сходятся на том, что семантических критериев явно недостаточно для понимания сложной природы грамматического рода и необходимо использовать все лингвистические факты, относящиеся к этой категории.

Попытаемся сначала объяснить (не строгим образом, а в расчете на интуитивное понимание), какова формальная природа соотношения между естественным и грамматическим родом. Возьмем в качестве исходного два существительных  $\xi$  и  $\eta$ . Хотим найти такое определение мужского и женского грамматических родов, которое бы в явной форме представило тот механизм, те операции, посредством которых эти категории получены из соответствующих естественных родов. Будем исходить из того, что любое существительное мужского рода должно находиться в строго определенном

формальном отношении с существительным  $\xi$ , и любое существительное женского рода должно находиться в подобном же отношении с существительным  $\eta$ . Чтобы получить такое отношение, обратимся к понятию цепи, введенному Ревзиным.

*Если специально не оговаривается противное, то во всех последующих примерах множество предложений будет состоять лишь из синтагм типа существительное + качественное прилагательное в положительной степени или качественное прилагательное в положительной степени + существительное.*

Понятия цепи и длины цепи позволяют получить такое определение мужского и женского грамматических родов, которое предполагает переход от естественного к грамматическому роду. Правила, представленные далее, объясняют многие конкретные факты естественных языков (на примере французского языка).

*Существительное относится к мужскому грамматическому роду, если любая словоформа его парадигмы может быть соединена с любой словоформой парадигмы слова  $\xi$  посредством цепи, длина которой не превышает 3.* Из этого непосредственно следует, что и  $\xi$  принадлежит мужскому грамматическому роду.

*Существительное относится к женскому грамматическому роду, если любая словоформа его парадигмы может быть соединена с любой словоформой парадигмы слова  $\eta$  посредством цепочки, длина которой не превышает 3.* Отсюда непосредственно следует, что  $\eta$  принадлежит женскому грамматическому роду.

*Существительное относится к среднему роду, если оно не относится ни к женскому, ни к мужскому грамматическому роду.*

*Существительное относится к общему роду, если оно принадлежит как мужскому, так и женскому роду.*

Эти правила характеризуют категорию грамматического рода во многих естественных языках. Считается, что эта категория является невырожденной для существительных данного языка в том случае, когда в этом языке существует, по крайней мере, одно существительное мужского рода, которое нельзя отнести к женскому роду, и, по крайней мере, одно существительное женского рода, которое нельзя отнести к мужскому роду. Рассмотрим приведенные выше правила на примере французского языка.

Пусть  $\xi = \text{homme}$ ,  $\eta = \text{femme}$ . Существительные *colins*, *cahier*, *murs* и т.д. относятся к мужскому роду, так как имеется цепь: (a) *colins*, *colin*,

*homme*, (b) *colins, hommes*, (c) *colin, homme, hommes*, (d) *cahier, homme*, (e) *cahiers, hommes, homme*, (f) *cahier, homme, hommes*, (g) *murs, mur, homme*, (h) *murs, hommes*, (i) *mur, homme, hommes*; существительные *plumes, pluie, feuille* и другие принадлежат женскому роду, так как любая словоформа из парадигм этих существительных может быть соединена с любой формой слова *femme* цепью, длина которой не больше 3; имеются цепи *plumes, plume, femme; pluie, femme, femmes; feuille, femme* и т.д. Существительные *cas, tas, tapis, nez, voix* и другие, у которых форма единственного числа совпадает с формой множественного числа, являются существительными среднего рода, так как нет ни одной формы слова *homme* и ни одной формы слова *femme*, которые принадлежали бы тому же дистрибутивному классу, что и *cas* (имеются формы *petit cas, petits cas*, но нет ни *petit hommes*, ни *petits homme*), так же как не существует никакой формы *homme* и никакой формы *femme*, которые принадлежали бы тому же дистрибутивному классу, что и *voix* (имеются формы *belle voix, belles voix*, но нет ни *belle femmes*, ни *belles femme*; имеются *homme beau, hommes beaux*, но нет ни *voix beau*, ни *voix beaux*). Существительные *camarade, élève, enfant* и другие, которым могут предшествовать как прилагательные в форме мужского рода, так и женского, относятся к среднему роду, так как дистрибутивные классы таких слов не совпадают с дистрибутивными классами ни словоформ *homme*, ни словоформ *femme*. Действительно, имеются *bon élève, bonne élève*, но нет ни *bon femme*, ни *bonne homme*; имеются *bons élèves, bonnes élèves*, но отсутствуют *bons femmes* и *bonnes hommes*.

Итак, во французском языке имеются существительные, которые относятся к мужскому роду и не относятся к женскому (*cahier, mur, soleil* и т.д.), а также существительные, которые относятся к женскому роду, но не относятся к мужскому (*plume, pluie, feuille* и т.д.); значит, грамматический род французских существительных не является вырожденным.

Сделаем следующие замечания.

1. Некоторые незначительные противоречия связаны с ограниченным запасом контекстов (прилагательное + существительное, существительное + прилагательное). Если соответствующим образом преобразовать исходное множество отмеченных последовательностей, то подобные противоречия исчезнут.

2. Приписывание словам грамматических родов не абсолютно однозначно. Оно зависит от определенного выбора парадигм и предложений (отмеченных последовательностей слов). Если мы, например, возьмем только изолированные слова в качестве предложений, то все они образуют единственный дистрибутивный класс, так что, взяв любые два слова, мы

можем соединить их цепью, длина которой равна 2. В частности, все существительные будут относиться как к мужскому, так и к женскому роду, и следовательно, грамматический род будет вырожденным. Это замечание раскрывает синтагматический характер грамматического рода, его существенную зависимость от контекста.

Пусть  $\{G, P, \Phi\}$  — некоторый язык. Будем говорить, что два слова  $a \in G$  и  $b \in G$  принадлежат одному и тому же роду, и будем обозначать это  $a \rho b$ , если для всякого  $a' \in P(a)$  и всякого  $b' \in P(b)$  выполняется, по крайней мере, одно из двух следующих условий:  $P(a) \cap S(b') \neq \emptyset$ ,  $P(b) \cap S(a') \neq \emptyset$ . Будем говорить, что  $a$  и  $b$  принадлежат одному и тому же ограниченному роду, и будем записывать это  $a \rho b$ , если для всякого  $a' \in P(a)$  и всякого  $b' \in P(b)$  верно, что  $P(a) \cap S(b') \neq \emptyset \neq P(b) \cap S(a')$ .

Становится ясно, что две словоформы могут иметь один и тот же род, даже если они не принадлежат одной и той же части речи. Из-за этого факта возникает противоречие между грамматическим родом и его математической моделью, потому что имеет смысл говорить о роде только для отдельных частей речи. Например, грамматический род у прилагательных совсем иной, чем у существительных; наши модели имеют отношение лишь к существительным.

Перейдем теперь к проблеме измерения различий между двумя данными родами. С этой целью введем понятие *расстояния между двумя родами*  $G(x)$  и  $G(y)$ , которое определяется, как наименьшее число  $n$  такое, что всякое слово рода  $G(x)$  может быть соединено со всяким словом рода  $G(y)$  цепью, длина которой не более  $n + 1$ . В том случае, когда такого числа не существует, мы будем говорить, что расстояние между  $G(x)$  и  $G(y)$  бесконечно. Будем обозначать через  $\delta(x, y)$  расстояние между  $G(x)$  и  $G(y)$ . Нетрудно видеть, что  $\delta(x, y) \geq 0$ ,  $\delta(x, y) = \delta(y, x)$ ,  $\delta(x, y) \leq \delta(x, z) + \delta(z, y)$  для всякого слова  $z$ . Это понятие особенно удобно, когда множества слов, относящихся к разным родам, не пересекаются. Рассмотрим понятие расстояния между родами на примере французского языка.

Рассмотрим семь следующих родов:  $G(\text{crayon})$  — первый мужской,  $G(\text{arbre})$  — второй мужской,  $G(\text{maison})$  — женский,  $G(\text{cas})$  — первый средний,  $G(\text{voix})$  — второй средний,  $G(\text{enfant})$  — третий средний и  $G(\text{camarade})$  — четвертый средний. Поскольку  $\text{bel arbre}$  и  $\text{bels arbres}$  — правильные слово-

сочетания в отличие от *bel crayon*, *bels crayons*, то  $\delta(\text{crayon}, \text{arbre}) = \infty$ . С другой стороны, так как *beau crayon* и *beaux crayons* — правильные словосочетания, а *beau maison* и *beaux maisons* таковыми не являются, то  $\delta(\text{crayon}, \text{maison}) = \infty$ . Словосочетания *nouveau cas* и *nouveaux cas* правильные, а *nouveaux crayon* и *nouveau crayon* неправильные, поэтому  $\delta(\text{crayon}, \text{cas}) = \infty$ . Поскольку *beau voix* и *beaux voix* неправильны, то  $\delta(\text{crayon}, \text{voix}) = \infty$ . Ввиду того что правильные словосочетания *belle enfant* и *belles enfants* становятся неправильными, если заменить *enfant* на *crayon* и *enfants* на *crayons*,  $\delta(\text{crayon}, \text{enfant}) = \infty$ . Поскольку *bel arbre*, *bels arbres* правильно, тогда как *bel cas* и *bels cas* неправильно, то  $\delta(\text{arbre}, \text{cas}) = \infty$ . Ввиду того что словосочетания *bel voix*, *bels voix* неправильны, то  $\delta(\text{arbre}, \text{voix}) = \infty$ . Ввиду того что *belle enfant* и *belles enfants* правильны, тогда как *belle arbre*, *belles arbres* правильными не являются,  $\delta(\text{arbre}, \text{enfant}) = \infty$ . Поскольку *grande voix* и *grandes voix* — правильные словосочетания, а *grande maisons* и *grandes maison* — нет, то  $\delta(\text{maison}, \text{voix}) = \infty$ . Поскольку *bel enfant* и *bels enfants* являются правильными словосочетаниями, тогда как *bel maison* и *bels maisons* — нет, то  $\delta(\text{maison}, \text{enfant}) = \infty$ . Подобным же образом мы находим, что  $\delta(\text{maison}, \text{cas}) = \infty$ ,  $\delta(\text{cas}, \text{voix}) = \infty$  и  $\delta(\text{voix}, \text{enfant}) = \infty$ . Из того, что словосочетания *beau camarade*, *beaux camarades*, *bell camarade* и *belles camarades* являются правильными, тогда как *beaux camarade*, *beau camarades*, *belles camarade*, *belle camarades*, *bel camarade*, *bels camarades* неправильны, следует, что имеют место равенства

$$\delta(\text{crayon}, \text{camarade}) = \delta(\text{arbre}, \text{camarade}) = \delta(\text{maison}, \text{camarade}) = \\ \delta(\text{cas}, \text{camarade}) = \delta(\text{voix}, \text{camarade}) = \delta(\text{enfant}, \text{camarade}) = \infty.$$

Нетрудно видеть, что два различных рода никогда не пересекаются.

#### 4.4. Категории падежа

Пусть  $\Gamma$  — фиксированное множество элементов, называемых словами,  $\mathcal{P}$  — разбиение множества  $\Gamma$  и  $\Phi$  — подмножество свободной полугруппы над  $\Gamma$ . Элементами этой полугруппы являются последовательности слов, а элементами  $\Phi$  — отмеченные или маркированные последовательности. Предполагается, что каждое слово входит хотя бы в одну отмеченную по-

следовательность. Если  $x \in \Gamma$ , то  $P(x)$  — это элемент разбиения  $P$ , содержащий  $x$ ;  $y \in P(x)$  будет пониматься как форма слова  $x$ .  $P(x)$  будет *парадигмой* слова  $x$ . Как и прежде, тройка  $\{\Gamma, P, \Phi\}$  будет называться *языком с парадигматической структурой*, или просто *парадигматическим языком*.

*Контекстом* мы будем называть упорядоченную пару последовательностей  $(f_1, f_2)$ , причем последовательности  $f_1$  и  $f_2$  не обязательно являются отмеченными. Контекст  $(f_1, f_2)$  называется *отмеченным*, если существует такое слово  $x$ , что последовательность  $f_1 x f_2$  является отмеченной.

Слово  $a$  является *допустимым* в контексте  $(f_1, f_2)$ , если существует такое слово  $a' \in P(a)$ , что последовательность  $f_1 a' f_2$  отмечена.

Слово  $a$  *непосредственно допустимо* в контексте  $(f_1, f_2)$ , если последовательность  $f_1 a f_2$  является отмеченной. Очевидно, что всякое слово, непосредственно допустимое в контексте  $(f_1, f_2)$ , является одновременно допустимым в  $(f_1, f_2)$ , однако обратное неверно. Обозначим через  $A(\gamma)$  множество слов, допустимых в контексте  $\gamma = (f_1, f_2)$ .

Два контекста  $\gamma'$  и  $\gamma''$  будем называть *эквивалентными*, если  $A(\gamma') = A(\gamma'')$ . Пусть  $B(\gamma)$  — множество слов, непосредственно допустимых в контексте  $\gamma$ . Очевидно, что  $B(\gamma) \subseteq A(\gamma)$ . Обратное не всегда верно.

Два контекста  $\gamma'$  и  $\gamma''$  *строго эквивалентны*, если  $B(\gamma') = B(\gamma'')$ . Нетрудно понять, что два строго эквивалентных контекста эквивалентны.

Пусть  $\mathbf{E}$  — подмножество  $\Gamma$ . Будем говорить, что  $\mathbf{E}$  — *допустимое множество*, если существует хотя бы один такой контекст  $\gamma$ , что  $\mathbf{E} = A(\gamma)$ .

Назовем  $\mathbf{E}$  *непосредственно допустимым*, если существует хотя бы один такой контекст  $\gamma$ , что  $\mathbf{E} = B(\gamma)$ . Будем говорить, что множества  $\mathbf{E}$  и  $\mathbf{H}$  являются *одновременно допустимыми*, если существует такой контекст  $\gamma$ , что  $\mathbf{E} = A(\gamma)$  и  $\mathbf{H} = B(\gamma)$ .

Мы будем говорить, что два контекста  $\gamma'$  и  $\gamma''$  *согласованы относительно*  $a \in \Gamma$ , если существует такое слово  $a' \in P(a)$ , что  $a'$  непосредственно допустимо одновременно и в  $\gamma'$ , и в  $\gamma''$ .

Мы будем говорить, что  $\gamma'$  и  $\gamma''$  *полностью согласованы относительно*  $a \in \Gamma$ , если  $P(a) \cap B(\gamma') = P(a) \cap B(\gamma'') \neq \emptyset$ .

Очевидно, что отношения согласования и полного согласования относительно слова  $\alpha$  симметричны на множестве контекстов языка.

Введем еще два новых понятия. Пусть даны два контекста  $\gamma_1$  и  $\gamma_2$ . Мы будем говорить, что  $\gamma_1$  и  $\gamma_2$  *конгруэнтны*, и будем обозначать это  $\gamma_1 \approx \gamma_2$ , если существует конечная цепочка контекстов  $\delta_1, \delta_2, \dots, \delta_i, \delta_{i+1}, \dots, \delta_n$ , обладающая тремя следующими свойствами: (1)  $\delta_1 = \gamma_1$ ; (2)  $\delta_n = \gamma_2$ ; (3) для любого  $i < n$  контексты  $\delta_i$  и  $\delta_{i+1}$  согласованы.

Если в третьем условии вышеприведенного определения заменить термин «согласованы» на «полностью согласованы», то мы получим новое отношение. В этом случае мы будем говорить, что  $\gamma_1$  и  $\gamma_2$  *полностью конгруэнтны*, и будем обозначать это  $\gamma_1 \approx \gamma_2$ .

Очевидно, что два согласованных контекста являются конгруэнтными и что два полностью согласованных контекста полностью конгруэнтны. Обратное утверждение неверно.

Отмеченные контексты образуют разбиение на классы конгруэнтности, т.е. на классы эквивалентности относительно  $\approx$ . Каждый класс конгруэнтности на множестве отмеченных контекстов мы будем называть *грамматическим падежом* или просто *падежом*.

Отсюда ясно, что каждый отмеченный контекст соотнесен с каким-либо грамматическим падежом, причем только с одним. Отмеченные контексты также образуют разбиение относительно полной конгруэнтности. Всякий класс полной конгруэнтности мы назовем совмещенным грамматическим падежом.

Тем не менее, можно отметить следующее несовершенство модели. Контексты (*я вижу синий,  $\theta$* ) и (*синий,  $\theta$* ) являются согласованными, а значит, и конгруэнтными, хотя с точки зрения традиционной грамматики мы имеем здесь разные падежи. Чтобы обойти эту трудность, можно отказаться от некоторых контекстов с прилагательными, порядковыми числительными и т.п.

К числу недостатков модели относится и то, что разные словоформы одного и того же существительного могут быть непосредственно допустимыми в одном контексте. Например, формы *газеты* и *газету* непосредственно допустимы в контексте (не читал,  $\theta$ ), формы *кошке* и *кошку* непосредственно допустимы в контексте (дал,  $\theta$ ). Чтобы избежать этого, аналогично ограничивают классы рассматриваемых контекстов таким образом, чтобы две различные словоформы одного и того же слова не могли быть непосредственно допустимыми в одном контексте.

## 5. СЛОВООБРАЗОВАТЕЛЬНЫЕ КОНСТРУКЦИИ

В данном разделе приведены основные словообразовательные конструкции русского языка, *суффиксация* от основ прилагательных и глаголов и другие методы. В частности, этими способами образуются большинство существительных.

### 5.1. Словообразование имен существительных

Для *префиксальных* существительных наиболее характерны следующие словообразовательные категории.

1. Производные с общим значением интенсивности, высокой степени того, что названо производящей основой:

*раз-*: красавица — раскрасавица, граф — разграф;

*сверх-*: человек — сверхчеловек, прибыль — сверхприбыль;

*супер-*: экспресс — суперэкспресс, кубок — суперкубок;

*архи-*: демократ — архидемократ, порядок — архипорядок;

*ультра-*: мода — ультрамода, звук — ультразвук;

*гипер-*: вежливость — гипервежливость, фонды — гиперфонды;

*экстра-*: мода — экстрамода, совершенство — экстрасовершенство.

2. Производные с общим значением противоположности, отрицания:

*анти-*: частица — античастица, герой — антигерой;

*противо-*: меры — противомеры, действие — противодействие;

*контр-*: удар — контрудар, позиция — контрпозиция, инициатива — контринициатива;

*де-/дез-*: гуманизация — дегуманизация; информация — дезинформация;

*не-*: счастье — несчастье, покой — непокой;

*дис-*: пропорция — диспропорция, гармония — дисгармония, комфорт — дискомфорт.

3. Производные с общим значением неистинности, ложности:

*псевдо-*: рецензия — псевдо рецензия, открытие — псевдооткрытие;

*квази-*: диалог — квазидиалог, открытие — квазиоткрытие.

4. Производные со значением совместности:

*со-*: автор — соавтор, переживание — сопереживание.

5. Производные со значением подчиненности:

*под-*: вид — подвид, группа — подгруппа, язык — подъязык;

*суб-*: продукты — субпродукты, инспектор — субинспектор, аренда — субаренда.

Кроме того, в словообразовании существительных действуют отдельные малопродуктивные приставки.

Префиксально-суффиксальным способом образуются существительные на базе сочетаний существительных с предлогами. В составе производного предлог преобразуется в приставку. Производящим является сочетание имени существительного с предлогом. Производные, как правило, обозначают предметы, соотнесенные в пространственном или временном отношении с тем, что названо производящей основой. Наиболее употребительны при этом суффиксы *-ник*, *-j(e)*, *-иц(a)*.

1. Производные с пространственными значениями:

*под-...-ник* (предмет, находящийся под тем, что названо производящей основой): *под снегом* — *подснежник*, *под локтем* — *подлокотник*;

*на-...-ник* (предмет, находящийся на том, что названо производящей основой): *на колене* — *наколенник*, *над почкой* — *надпочечник*, *на конце* — *наконечник*;

*за-...-j(e)* (то, что находится за предметом, названным производящей основой): *за рекой* — *заречье*, *за Волгой* — *Заволжье*;

*при-...-j(e)* (то, что находится рядом с предметом, названным производящей основой): *при Урале* — *Приуралье*, *при море* — *приморье*;

*меж-...-j(e)* (то, что находится между предметами, названными производящей основой): *меж горами* — *межгорье*, *меж бровями* — *межбровье*.

2. Производные с временными значениями:

*меж-...-j(e)*: *меж сезонами* — *межсезонье*;

*пред-...-j(e)*: *пред зимой* — *предзимье*.

3. Производные, обозначающие отсутствие того, что названо производящей основой, образуются с помощью приставки *без-* и суффиксов *-иц(a)* и *-j(e)*:

*без-...-иц(a)*: *без работы* — *безработица*;

*без-...-j(e)*: *без денег* — *безденежье*, *без людей* — *безлюдье*, *без земли* — *безземелье*, *без культуры* — *бескультурье*.

Группа префиксально-суффиксальных существительных, которая не соотносится с сочетанием предлог + существительное. Приставка, входящая в

производное, не имеет омонимичного предлога. Эта группа малочисленна и дает мало новообразований. В производных участвуют приставка *анти-* и суффикс *-ин*: *комар* — *антикомарин*, шуточные: *антиведмин*, *антилюбовин*. Производные называют средство против того, что названо основой производящего существительного.

Префиксация в соединении с нулевой суффиксацией используются для производства от основ имен прилагательных существительных со значением «слабой степени обнаружения какого-либо признака, названного производящей основой». При этом действует префикс *про-* в соединении с нулевым суффиксом.

Производные содержат те же чередования, что слова типа *сушь*, и относятся к тому же склонению: *желтый* — *прожелть*, *седой* — *просесть*.

Образование сложных существительных происходит с помощью сложения, а также слиянием способа сложения с префиксальным, суффиксальным и префиксально-суффиксальным способами.

## 5.2. Словообразование имен прилагательных

В словообразовании имен прилагательных развиты суффиксация, префиксация, префиксально-суффиксальный способ, словосложение и сращение.

Префиксация — один из самых распространенных способов словообразования прилагательных.

Наиболее многочисленна категория прилагательных, обозначающих интенсивность, полноту проявления признака:

**наи-** (прилагательное обычно содержит суффикс *-ейш-/-айш-*): *умнейший* — *наиумнейший*, *глупейший* — *наиглупейший*;

**пре-**: *умный* — *преумный*, *глупый* — *преглупый*; чаще в повторах: *добрый-предобрый*, *милый-премилый*;

**раз-**: *веселый* — *развеселый*; чаще в повторах: *удалый-разудалый*;

**пере-** (обычно в повторах): *тертый-перетертый*, *писанный-переписанный*;

**архи-**: *прозаический* — *архипрозаический*, *благородный* — *архиблагородный*;

**сверх-**: *молодой* — *сверхмолодой*, *одаренный* — *сверходаренный*;

**супер-**: *модный* — *супермодный*, *тяжелый* — *супертяжелый*;

**экстра-**: *новый* — *экстрановый*, *модный* — *экстрамодный*;

**ультра-**: *оригинальный* — *ультраоригинальный*, *современный* — *ультрасовременный*.

Другую словообразовательную категорию составляют прилагательные, в которых приставки имеют значение отрицания, противоположности:

**не-**: умный — неумный, красивый — некрасивый;

**анти-**: гуманный — антигуманный, военный — антивоенный;

**противо-**: воспалительный — противовоспалительный («направленный против воспаления»), химический — противохимический, атомный — противоатомный («направленный против химического или атомного оружия»).

Префиксально-суффиксальным способом образуются прилагательные на базе именных и глагольных сочетаний.

Прилагательные, производимые на базе сочетаний существительных с предлогами, преобразуемыми в составе производных в приставки, могут включать разнообразные приставки, а из суффиксов *-н-* (преимущественно), реже *-ов-*, *-ск-* и некоторые другие:

**без-**: без пилота — беспилотный, без шума — бесшумный;

**вне-**: вне завода — внезаводской, вне земли — внеземельный;

**внутри-**: внутри института — внутриинститутский, внутри школы — внутришкольный;

**до-**: до школы — дошкольный, до посева — допосевный;

**за-**: за бортом — забортовой, за рекой — заречный;

**меж-**: меж планетами — межпланетный, меж районами — межрайонный;

**на-**: на горе — нагорный, на столе — настольный;

**над-**: над глазом — надглазный, над окном — надоконный;

**около-**: около звезды — околозвездный, около солнца — околосолнечный;

**под-**: под вагоном — подвагонный;

**после-**: после полета — послеполетный, после посева — послепосевный;

**перед-**: перед съездом — предсъездовский, перед дипломом — преддипломный;

**противо-**: против пожара — противопожарный;

**через-**: через полосу — чересполосный.

На базе соединения отрицания *не* с сочетанием имени существительного с предлогом *без* образуются имена прилагательные со сложной приставкой *небез-* и суффиксом *-н-*. Производное имеет значение неполноты, слабой степени проявления признака: *не без пользы* — *небесполезный*, *не без греха* — *небезгрешный*, *не без вреда* — *небезвредный*. Прилагательные с приставкой *небез-* отличаются по значению от прилагательных с приставкой *без-* и

отрицанием *не-*. Например: *Этот напиток не безвреден* (то есть вреден, отрицается безвредность напитка). — *Этот напиток небезвреден* (то есть вреден в слабой степени).

Прилагательные, образованные префиксально-суффиксальным способом от основ глаголов, образуются с помощью приставки *не-* и суффиксов *-н-* и *-м-*. Производные обозначают «невозможность подвергнуться действию»:

***не-...-н(ый)***: *разорвать* — *неразрывный*, *возвратить* — *невозвратный*;

***не-...-м(ый)***: *исцелить* — *неисцелимый*, *переносить* — *непереносимый*.

Префиксация в соединении с нулевой суффиксацией используется для производства прилагательных на базе сочетаний предлога *без* с именами существительными. Обычно эти существительные называют часть тела человека или животного. Производное обозначает «не имеющий того, что названо производящей основой»: *без ноги* — *безногий*, *без руки* — *безрукий*. Реже производящая основа относится к иному семантическому разряду: *без листьев* — *безлистный*, *без коры* — *безкорый*. Возможны однокорневые образования с тем же значением, содержащие с своим составе суффикс: *безлистый* — *безлистный*, *безъязыкий* — *безъязычный*.

### 5.3. Словообразование глаголов

В словообразовании глаголов главенствующее место занимает внутриглагольное словообразование, т. е. производство от основ глагола чистой префиксацией и префиксально-суффиксальным способом. Суффиксация имеет меньшее значение, а словосложение для словообразования глагола совсем мало характерно.

Среди префиксальных глаголов различаются две большие группы: с приставками пространственных значений и с приставками количественно-временных значений.

Глаголы с приставками пространственных значений обозначают различные направления действия:

***в-***: *бежать* — *вбежать*, *нести* — *внести*;

***вз-***: *бежать* — *взбежать* (в гору), *лететь* — *взлететь*;

***до-***: *бежать* — *добежать*, *лететь* — *долететь*;

***за-***: *бежать* — *забежать* (за что-нибудь), *лететь* — *залететь* (за что-нибудь, куда-нибудь);

***на-***: *бежать* — *набежать*, *лететь* — *налететь*;

***о-/об-***: *бежать* — *обежать*, *лететь* — *облететь* (что-нибудь);

***от-***: *бежать* — *отбежать*, *лететь* — *отлететь*;

**пере-**: бежать — перебежать, лететь — перелететь (через реку);

**при-**: бежать — прибежать, лететь — прилететь;

**про-**: бежать — пробежать (под мостом), лезть — пролезть;

**с-**: бежать — сбежать (с горы), лезть — слезть (со стола);

**у-**: бежать — убежать, лететь — улететь.

Глаголы с приставками количественно-временных значений обозначают временные пределы, а также силу, интенсивность или слабость, неполноту действия.

1. Глаголы, обозначающие начало процесса:

**за-**: петь — запеть, говорить — заговорить;

**по-**: бежать — побежать, плыть — поплыть;

**вз-** (с оттенком интенсивности действия): кричать — вскричать, реветь — взрывать.

2. Глаголы, обозначающие окончание процесса:

**от-** («прекращение процесса»): цвести — отцвести, греметь — отгреметь;

**до-** («доведение действия до конца»): думать — додумать, читать — дочитать, шить — дошить.

3. Глаголы, обозначающие окончание действия с оттенками полноты, тщательности, энергичности, силы его выполнения:

**раз-**: ругать — разругать, баловать — разбаловать, целовать — расцеловать;

**пере-**: жарить — пережарить, кормить — перекормить;

**на-**: безобразничать — набезобразничать, возить — навозить;

**про-**: жарить — прожарить, варить — проварить (как следует), сушить — просушить;

**вы-**: мерить — вымерить, строгать — выстрогать.

4. Глаголы, обозначающие полную исчерпанность предмета действием, а также причинение неприятности, ущерба действием:

**вы-**: топтать — вытоптать, рубить — вырубить;

**из-**: мылить — измылить (все мыло), писать — исписать (карандаш);

**за-**: сечи — засечь (до смерти), играть — заиграть (мелодию);

**с-**: носить — сносить (башмаки), винтить — свинтить (винт);

**от-**: сидеть — отсидеть (руку), топтать — оттоптать (ноги);

**про-**: носить — проносить (брюки до дырки), плакать — проплакать (все глаза).

5. Глаголы, обозначающие дополнительное, добавочное действие, добавление чего-либо действием, а также слабость, неполноту действия:

**до-** («дополнительное действие»): *грузить* — *догрузить*, *купить* — *докупить*;

**под-** («дополнительное действие — для количественного увеличения объекта», «прибавление действием»): *лить* — *подлить*, *купить* — *подкупить*;

**под-** («дополнительное, повторное действие — для более полного достижения результата»): *варить* — *подварить* (кашу; = доварить), *красить* — *подкрасить*;

**при-** (неполнота действия): *встать* — *привстать*, *лечь* — *прилечь*;

**над-**: *рубить* — *надрубить*, *колоть* — *надколоть*, *бить* — *надбить*;

**недо-**: *выполнить* — *недовыполнить*, *думать* — *недодумать*, *рассказать* — *недорассказать*.

Префиксально-суффиксальным способом образуются глаголы от основ глаголов и от основ других частей речи.

Глаголы, производимые от основ глаголов, образуют две группы:

а) приставка + производящая основа + суффикс несовершенного вида;

б) приставка + производящий глагол + —ся.

Рассмотрим каждую группу более детально.

1. Неполноту, ослабленность действия обозначает группа одновидовых глаголов несовершенного вида, имеющих прерывисто-длительное значение:

**по-**: *трещать* — *потрескивать*, *кричать* — *покрикивать*;

**при-** (с оттенком дополнительности действия): *плясать* — *приплясывать*, *говорить* — *приговаривать*;

**под-** (с оттенком дополнительности действия): *петь* — *подпевать*, *свистеть* — *подсвистывать*;

**на-**: *петь* — *напевать*, *свистеть* — *насвистывать*, *шептать* — *нашептывать*.

Интенсивность, тщательность совершения действия выражает группа глаголов с приставками:

**вы-**: *шагать* — *вышагивать*, *плясать* — *выплясывать*;

**на-**: *звонить* — *названивать*, *хвалить* — *нахвалять*.

2. Глаголы, образуемые одновременным присоединением приставки и -ся, многообразны по значениям:

**в-...-ся** («предельная исчерпанность действия, углубленность в действие»): *читать* — *вчитаться*, *думать* — *вдуматься*;

**за-...-ся** («увлеченность действием»): *думать* — *задуматься*, *читать* — *зачитаться*;

**на-...-ся** («удовлетворенность действием»): *гулять* — *нагуляться*, *работать* — *наработаться*;

**из-...-ся** («исчерпанность в результате действия, приобретение каких-либо отрицательных свойств в результате действия»): *хулиганить* — *исхулиганиться*, *писать* — *истисаться*;

**до-...-ся** («достижение результата после интенсивного выполнения действия»): *звать* — *дозваться*, *звонить* — *дозвониться*.

Глаголы, производимые от основ имен существительных префиксально-суффиксальным способом, используют различные приставки в соединении с суффиксами -Ø -/-(и)ть и -е(ть):

**за-...-ить**: *штора* — *зашторить*;

**об-...-ить**: *лес* — *облесить*;

**при-...-ить**: *земля* — *приземлить*.

Для отыменного производства глаголов префиксально-суффиксальным способом характерен префикс *обез-*. Производные на *-ить* обозначают «лишить того, что названо производящей основой»: *вред* — *обезвредить*. Производные на *-еть* обозначают «лишиться того, что названо производящей основой»: *сила* — *обессилеть*, *ум* — *обезуметь*.

Глаголы, производимые от основ имен прилагательных суффиксально-префиксальным способом, производятся с помощью различных приставок (чаще всего *у-* и *о-*) и тех же суффиксов -Ø -/-(и)ть и -е(ть): *плотный* — *уплотнить*, *глубокий* — *углубить*, *благородный* — *облагородить*, *слабый* — *ослабеть*, *скудный* — *оскудеть*.

#### 5.4. Наречие как неизменяемая часть речи и типы наречий

Наречия — это такие неизменяемые полнозначные слова, которые способны называть признак действия или признак признака. Есть наречия, способные называть признак предмета: *яйцо всмятку*, *разговор по-английски* и наречия, называющие признак предмета или признак признака действия: *очень красивое лицо*, *или очень долго*. Наречия не могут определяться прилагательными.

##### 5.4.1. Семантические разряды наречий

Из определения наречий следует, что их синтаксическая сочетаемость весьма разнообразна. С различием в синтаксической сочетаемости связано и различие в семантических разрядах. Наречия со значением меры и степени (*много*, *очень*, *совсем*, *чуть-чуть*) в принципе могут относиться к глагольным формам, к имени прилагательному и наречию. Однако при наличии семантической общности каждое наречие из этой группы обладает собст-

венными сочетательными свойствами (ср.: *много говорить* — *очень разговорчивый*; *едва видеть* — *едва видный*).

Наречия также могут иметь значение качественной характеристики действия: *громко кричать*, *чисто писать*; образа действия: *ехать верхом*, *смотреть исподлобья*; времени: *прийти вовремя*, *сказать завтра*; места: *смотреть кругом*, *оглянуться назад*; причины: *взять сослепу*, *сказать сгоряча*; цели: *сделать назло*, *написать нарочно*, *сказать умышленно*. Наречия данной семантики сочетаются обычно с глагольными формами, однако не полностью исключена и их сочетаемость с существительными (ср.: *взглянуть вперёд* — *взгляд вперёд*).

Практически полезным является правило: если при проверяемом слове может быть определяющее его прилагательное — перед нами существительное; если появление прилагательного невозможно — перед нами наречие. Например, *шёл берегом*, но *шёл высоким, красивым левым берегом*, следовательно, *берегом* — падежная форма существительного; *волосы ёжиком* — вставка прилагательного практически невозможна, следовательно, *ёжиком* — наречие.

Среди наречий, как и среди имен, есть местоименные слова: *там*, *потом*, *когда*, *так*, *зачем* и неместоименные: *медленно*, *громко*, *назло* и др.

#### 5.4.2. Степени сравнения наречий

Многие наречия имеют синтетические образования со значением большой степени качества: *смело* — *смелее (смелей)*, *внимательно* — *внимательнее (внимательней или внимательнейше)*. Степени сравнения наречий создают разные лексемы.

Сравнительная степенность прилагательных и наречий совпадает по форме и разграничить эти образования можно лишь по синтаксическим признакам. Если сравнительная степень характеризует действие, это — сравнительная степень наречия. Если сравнительная степень характеризует предмет, это — сравнительная степень прилагательного. Например: *Петя пишет красивее, чем Ваня* — наречие; *Петя красивее Вани* — прилагательное.

#### 5.4.2. Категория состояния и слова различных групп в синтаксической функции сказуемого

Состав категории состояния довольно широк. Это происходит, во-первых, из-за рассмотрения слов типа *скучно*, *весело*, *тоскливо*, употребляемых в функции сказуемого как категории состояния. В итоге получается, что в слове типа *скучно* мы имеем три омонима, принадлежащих к разным

частям речи: *Это занятие скучно* — краткое прилагательное; *Он рассказы-вал долго и скучно* — наречие; *Мне скучно* — категория состояния.

Во-вторых, категория состояния включает в себя такие употребляемые в функции сказуемого слова, как *грех, срам, пора, время, лень*. Таким образом, слово *грех* в предложении *Грех — не беда, молва нехороша*, является существительным, а в следующем предложении : *Над старостью смеяться грех* — категорией состояния.

Можно сказать, что категория состояния — это такие слова, которые, не будучи глаголами, выступают в функции сказуемого. При этом уже не обращается внимание на то, являются ли эти слова изменяемыми, выполняют ли они только функцию сказуемого.

#### 5.4.3. Объем группы слов, принадлежащих к категории состояния

К категории состояния, как правило, относят лишь неизменяемые слова. Все изменяемые слова — существительные, прилагательные, — даже употребленные в функции сказуемого, остаются соответственно существительными и прилагательными.

Сложнее обстоит дело со словами *пора, лень, грех, срам*. Когда эти слова выполняют функцию сказуемого, они меняют свою родовую характеристику, согласуясь с формами глагола прошедшего времени среднего рода. Ср.: *Ужасная пора (лень); Ужасный грех (срам) и Пора (лень, грех, срам) было так поступать*. Это обстоятельство — аргумент в пользу признания слов *пора, лень, грех, срам*, употребляемых в функции сказуемого, словами категории состояния.

Среди полнозначных неизменяемых слов можно выделить:

- 1) слова, способные выполнять функцию определителя признака или действия: *очень, быстро, неожиданно*;
- 2) слова, способные выполнять указанную функцию и функцию сказуемого: *светло, грустно, весело*;
- 3) слова, способные выполнять функцию сказуемого: *нужно, можно, жаль*.

К категории состояния следует относить только слова третьей группы. Рассматривая вторую группу, важно всякий раз оценивать, в какой именно синтаксической функции выступает слово. Однако из различия в этой функции не следует делать прямолинейный вывод о различиях в принадлежности к частям речи. Следует иметь в виду, что многие слова второй группы выступают преимущественно в функции сказуемого.

Итак, категорию состояния как часть речи составляют неизменяемые полнозначные слова, единственная синтаксическая функция которых — функция сказуемого.

#### 5.4.4. Модальные слова

Среди самостоятельных неизменяемых слов есть такие, которые вообще не бывают каким-либо членом предложения, но способны образовать слова-предложения: *вероятно, конечно, видимо* и т. п. Эти слова нельзя рассматривать ни как существительные, ни как наречия или категорию состояния.

Кроме синтаксической общности, самостоятельности и неизменяемости, данные слова обладают и общностью номинативной. Номинативная общность состоит в том, что с помощью этих слов выражается отношение говорящего к действительности, а также к форме и содержанию высказывания об этой действительности. Таким образом, в составе модальных слов оказываются не только слова типа *вероятно*, но и такие слова, как *во-первых, так, следовательно*.

Различают несколько разрядов модальных слов по их значению и употреблению:

- 1) слова, указывающие на чужие мысли, слова, стиль или их оценку: *буквально, так сказать*;
- 2) слова, указывающие на эмоциональную оценку действительности: *спасибо, полно, пожалуй*;
- 3) слова, указывающие на достоверность высказывания: *несомненно, безусловно, очевидно, видимо, в самом деле, подлинно*;
- 4) слова, указывающие на последовательность или характер связи мыслей или событий: *кроме того, в частности, в конце концов, во-первых, кстати*.

Исходя из приведенных определений наречия и категории состояния, представляется логичным рассматривать в качестве особой части речи неизменяемые знаменательные слова — модальные слова, не способные быть каким-либо членом предложения. При этом особенно важно отличать модальные слова как часть речи от других частей речи, употребленных в качестве вводных слов. Например, в предложениях *К сожалению примешивалось чувство облегчения* и *К сожалению, поездка не состоится* мы имеем предложно-падежную форму существительного как дополнение и как вводное слово, но перехода из одной части речи в другую не происходит. То же самое наблюдается и в предложениях *Он пришёл кстати* и *Кстати, я получил вчера письмо*.

### 5.5. Служебные части речи

Служебные части речи противопоставляются самостоятельным знаменательным (по-другому, — полнозначным) частям речи на основании принципа «функция в предложении».

Те слова, которые являются членами предложения и/или могут быть употреблены самостоятельно в качестве слова-предложения, формируют самостоятельные части речи. Слова, лишённые обеих таких возможностей, принадлежат к служебным частям речи.

Так, полнозначные слова *зима, молчу, хорошо* могут быть членами предложения и могут образовать предложения. Модальные слова типа: *безусловно, во-вторых, пожалуйста*, могут образовать предложения, однако не могут быть членами предложения. Относимые обычно к модальным словам языковые единицы типа *мол, дескать*, не могут быть членами предложения и не могут образовать предложения.

Служебные части речи делят на две группы по принципу «наличие или отсутствие функции выразителя связи между членами предложения или предложениями». Одни служебные слова связывают части предложения: *под, к, с, а, так как*. Другие служебные слова непосредственно связующих функций не выполняют: *бы, вот*.

Выполнение служебными словами связующей функции никак не означает отсутствия у них какого-либо номинативного значения. Указывая на наличие связи, служебные части речи одновременно определяют, каковы именно содержательные отношения между связываемыми единицами. Например, в словосочетаниях *шёл в метро, шёл от метро, шёл к метро* служебные слова не только указывают на то, что словоформы *шёл* и *метро* связаны между собой, но и несут в каждом случае определённую конкретную информацию о направлении действия относительно предмета. Ср. также: *видеть и знать — видеть, чтобы знать; не работал, но устал — не работал, потому что устал*.

Служебные части речи несколько напоминают окончания, поскольку многие окончания также одновременно и связывают слова в предложении, и отражают определённый аспект действительности.

Все служебные части речи являются неизменяемыми, у них отсутствуют какие-либо морфологические характеристики. Поэтому классификация внутри служебных частей речи может быть проведена только на основании принципов семантических и синтаксических.

### 5.5.1. Союз как служебная часть речи

Союз — это такая часть речи, к которой относятся слова, используемые для связи словоформ в простом предложении и частей сложного предложения, либо только для связи частей сложного предложения.

Например, союз *и* может связывать и словоформы в предложении: *Отец и сын*, и части сложного предложения: *Прозрачный лес один чернеет, и ель сквозь иней зеленеет*, а союз *чтобы* (*чтоб*) может соединять только части предложения: *Он рыбачил тридцать лет и три года и не слыхивал, чтоб рыба говорила*.

Союзы принято делить на непроеизводные: *а, но, и, как* и производные: *чтобы, оттого, потому что, в то время как* и др. Производные союзы могут быть простые: *чтобы, оттого* и составные: *после того, как; вследствие того, что; для того, чтобы; несмотря на то, что; как только*.

Составные союзы, на первый взгляд, выступают как словосочетания. Во всяком случае, они пишутся в виде нескольких слов и представляют собой такую последовательность, части которой похожи на самостоятельные лексемы. Однако данные языковые единицы не обладают релевантными признаками словосочетаний. «Части» рассматриваемых единиц имеют строго фиксированный порядок следования — вставить что-либо или невозможно: *потому что, в то время как*, или, наоборот, совершенно необходимо: *не только—но и, если—то, то...то*. Выделяются составные союзы, части которых совершенно различны: *не только—но и; если—то*, и союзы, части которых совпадают: *то-то, ни-ни* (их не следует путать с повторяющимися простыми союзами). Ср.: *А за окном **то** дождь, **то** снег; И пращ, и стрела, и лукавый кинжал щадят победителя годы* — в первом примере наличие первого *то* требует и второго *то*; во втором примере появление любого *и* теоретически совершенно независимо от наличия или отсутствия других *и*.

### 5.5.2. Классификация союзов

Как следует из определения союзов, среди них выделяются слова, которые соединяют однородные члены предложения, части сложносочиненных предложений и самостоятельные предложения. Такие союзы называют сочинительными: *и, а, но, или* и др.

Есть и такие слова, которые соединяют главную и придаточную части сложноподчиненного предложения. Такие союзы называют подчинительными: *после того как, ввиду того что, хотя* и др.

Как исключение в простом предложении подчинительные союзы могут связывать словоформы, которые не являются однородными членами пред-

ложения. Таков, например, союз *как*, присоединяющий сравнительный оборот или употребляющийся в значении «в качестве»: *Как невесту, родину мы любим; В городе он известен как хороший врач*. Другим исключением такого же типа является союз *чем* в составе оборота, включающего сравнительную степень: *Брат старше, чем сестра*.

Семантическая классификация союзов частично «перекрещивается» с их синтаксической классификацией. Сочинительные союзы по значению делят на соединительные, противительные, разделительные, градационные, пояснительные.

Соединительные союзы выражают отношения перечисления: *Пришёл, и ослабел, и лёг под сводом шалаша на лыки*; противительные — отношения различия, несоответствия: *Поклажа бы для них казалась и легка, да лебедь рвётся в облака, рак пятится назад, а щука тянет в воду*; разделительные — отношения взаимоисключения: *Или пан, или пропал*; градационные, называя явления, выделяют одно из них как более важное: *Он не только читает, но и хорошо говорит по-французски* (ср.: *Он по-французски совершенно мог изъясняться и писал*); пояснительные (*а именно, как-то, или*) указывают на различие в именовании того же самого явления.

Подчинительные союзы по значению делят на семантические, т. е. передающие определенные смысловые отношения, и асемантические, т. е. чисто синтаксические, указывающие лишь на «подчиненность» придаточной части сложноподчиненного предложения (*как, что, чтобы*). Семантические подчинительные союзы могут иметь значение времени: *когда, пока, едва, лишь, только, после того как* и др.; условия: *если, раз, ежели* и др.; причины: *потому что, так как, ибо* и др.; уступки: *хотя; несмотря на то, что* и др.; цели: *чтобы, затем чтобы* и др.; следствия: *так что* и сравнения: *как, словно, точно*.

### 5.5.3. Предлог как служебная часть речи

Предлог — это такая служебная часть речи, которая выражает различные «подчиненные» отношения существительного к другим словам в словосочетании и предложении. Предлоги, как и другие служебные части речи, принадлежат к неизменяемым словам.

По степени словообразовательной сложности выделяют предлоги непроизводные: *в, до, за, на, от, по, при, с* и производные: *вокруг, благодаря, в течение, навстречу* и др.

По признаку морфологической сложности среди производных предлогов следует выделить простые предлоги: *путём, спустя* и составные: *в связи с, независимо от, судя по* и др. Несмотря на особенности правописания,

составные предлоги представляют собой последовательности с жестким порядком следования частей.

Предлоги могут выражать различного рода отношения:

- пространственные (*к, от, под, над, за, в, из, на* и др.);
- временные (*после, в течение* и др.);
- причинные (*из-за, по причине* и др.).

Реже встречаются некоторые другие отношения, выражаемые предлогами. Как и союзы, предлоги могут выступать в чисто синтаксической функции. Например, в словосочетаниях *смеяться над бедой, верить в успех, надеяться на лучшее, бороться с недостатками* предлоги лишены какой-либо семантической функции, их присутствие обусловлено исключительно синтаксическими свойствами глаголов.

Многие предлоги многозначны. Эта многозначность, как и у других частей речи, устраняется благодаря контексту. Например: *шел из дома* — пространственное значение, *делал из зависти* — причинное.

Предлоги требуют, чтобы «подчиненное» им существительное было употреблено в определенной падежной форме. Поэтому предлоги делятся на такие, после которых непременно следует один падеж, либо два или даже три.

Р.п.	Д.п.	В.п.	Т.п.	П.п.
без	к	про	над	при
от	благодаря	сквозь	перед	
около	согласно	через		
из-за	вопреки			
	по			

Предлог *между* требует существительного в родительном или творительном падежах, предлог *в* — в винительном и предложном, предлог *под* — в винительном и творительном.

#### 5.5.4. Частицы

Частицы — это такие служебные слова, которые выражают смысловые и модально-экспрессивные оттенки предложений и слов и участвуют в образовании форм слова. Как и другие служебные слова, частицы не имеют форм словоизменения.

В настоящее время по признаку «функция» принято выделять частицы, участвующие в выражении грамматических значений наклонения совмест-

но с другими элементами глагольной формы (речь идет о сослагательном и повелительном наклонениях глагола и соответственно о частицах *бы, да, давай, пусть, пускай*), и частицы, имеющие лексические значения.

Частицы, имеющие лексические значения, подразделяются на несколько семантических разрядов: отрицательные частицы *не* и *ни*. Среди субъективно-модальных значений частиц выделяют усилительные: *Даже он пришёл; Он же знал об этом*; выделительные: *Только он пришел; Лишь он знал об этом*; вопросительные: *Неужели он пришёл?; Разве он знал об этом?*; восклицательные: *Ведь он пришёл!; Он куда как хорошо знал об этом?*; указательные: *вот, вон, это*; определительно-уточняющие: *именно, точно, приблизительно, почти*; утвердительные: *действительно, конечно, точно*; выражающие передачу чужой мысли: *мол, дескать* и некоторые другие. В этой классификации нет четкой границы между модальными (полнозначными) словами и частицами, поскольку некоторые слова, упомянутые выше как частицы (служебные слова), могут образовать слова-предложения и, следовательно, могут рассматриваться как модальные слова.

#### 5.5.5. Междометия и другие слова вне частей речи

Среди неизменяемых лексем особое место занимают междометия, слова-предложения, передающие, хотя и не называющие конкретно, различные чувства, такие, например, как удивление, испуг, гнев, радость: *Ба, знакомые всё лица; Ага, попались!*

Сам характер эмоций подчас не зависит прямо от лексического значения междометия, а возникает в результате контекстуальных воздействий. Например, междометие *ах* вне контекста может означать и радость, и испуг, и разочарование. Неизменяемость является чрезвычайно существенным свойством междометий. Изменяемые лексеммы, обладающие теми же семантическими свойствами, уже не междометия. Так, в предложении *Он никак не мог привыкнуть ко всем этим охам и ахам* слова «ох» и «ах» — существительные. Междометиями являются и сложные (состоящие не из одного корня) образования. Например, *боже мой, чёрт возьми* — междометия, каждое из которых представляет собой одну лексему, хотя графически в каждой из них выделяются две словоформы.

К междометиям относят и безаффиксные глагольные образования типа: прыг, толк: *Татьяна прыг в другие сени; Мартышка... тихохонько медведя толк ногой*. Строго говоря, у слов типов *ах, ба* и *прыг* мало общего; с междометиями их объединяет морфологическая общность, а именно неизменяемость.

К междометиям иногда относят и звукоподражательные слова типа *мяу-мяу*, *кис-кис*. Эти слова, не обозначая эмоций, по своим семантическим свойствам явно отличаются от слов типа *ах*. По семантическим и синтаксическим свойствам они противостоят и словам типа *прыг*. Считают, что звукоподражательные слова являются словами вне частей речи.

## 5.6. Местоименные слова

### 5.6.1. Категория рода и одушевленности—неодушевленности у местоименных существительных

К местоименным существительным относят местоимения личные: *я, ты, мы, вы, он, она, оно*; возвратные: *себя*; вопросительные: *кто, что*; неопределенные: *кто-то, что-то, кто-нибудь, что-нибудь, кто-либо, что-либо, кое-кто, кое-что, некто, нечто*; отрицательные: *никто, ничто, некого, нечего*.

Хотя местоименные существительные редко сочетаются с прилагательными, только по их сочетанию, как известно, можно определить родовую принадлежность существительного.

В соответствии с этим критерием местоименные существительные *он, кто-то, кто-нибудь, кто-либо, кое-кто, некто* принадлежат к мужскому роду; *она* — к женскому; *оно, что, что-то, что-нибудь, что-либо, кое-что, нечто* и *ничто* — к среднему.

Местоименные существительные *я, ты* принадлежат к общему роду: *я говорил (говорила), ты сказал (сказала)*. Местоимение *кто* может сочетаться с формами прилагательных мужского и женского рода, указывая соответственно на лиц мужского и (или) женского пола. Однако при обозначении лиц обоего пола местоимение *кто* сочетается с глаголом в формах прошедшего времени и сослагательного наклонения лишь в форме мужского рода: *Кто вышел замуж?*

Местоименное существительное *себя* может, видимо, иметь не одну родовую форму; ср.: *себя, больного; себя, больную*. Местоименные существительные *некого, нечего* можно также причислять и к мужскому и к среднему роду.

Все местоименные существительные обладают категорией одушевленности—неодушевленности: *я, ты, мы, вы, кто* и производные от него — одушевленные; *что* и производные от него — неодушевленные. Местоименные существительные *он, она, оно, они, себя* могут выступать и как одушевленные, и как неодушевленные: (*вижу*) *их всех — их все*. Одушев-

ленность или неодушевленность местоименных существительных зависит в этом случае от одушевленности или неодушевленности заменяемых местоимениями существительных: *вижу их всех — альпинистов, но вижу их все — вершины.*

### 5.6.2. Категория числа у местоименных существительных

Противопоставление по числу внутри парадигмы имеют местоименные существительные *он, она, оно*. Формы множественного числа у этих лексем омонимичны.

Личные местоименные существительные *я, ты и мы, вы* нельзя рассматривать как связанные только противопоставлением по числу. Ведь *мы* — это не «много я», но «я и другие»; *вы* — это не «много ты», а «совокупность различных лиц».

Остальные местоименные существительные по своим семантическим свойствам не чужды идее счета, но не обладают морфологическими средствами для выражения числового противопоставления. Они напоминают существительные типа *пальто и сани*. При этом местоименные существительные *себя, кто, кто-то, кто-нибудь, кто-либо, кое-кто* в принципе более или менее естественно могут выражать числовые противопоставления согласуемыми словоформами: *себя, больного — себя, больных; кто такой — кто такие* (ср.: *новое пальто — новые пальто*).

Другие не чуждые идее счета местоименные существительные не способны выражать противопоставление по числу и самой своей формой и согласуемыми словоформами: *что такое* возможно, а *что такие* — невозможно (ср.: *новые сани* — и об одних санях и о нескольких санях).

### 5.6.3. Категория падежа у местоименных существительных

У некоторых местоименных существительных нет части падежных форм. Не существует формы именительного падежа у местоименных существительных *себя, некого, нечего*.

Местоименное существительное *некто*, наоборот, имеет лишь эту форму. Только форму именительного-винительного падежа имеет *ничто*. В тех случаях, когда "предметы", обозначенные этими местоимениями, оказываются в контекстах, требующих иных падежных форм, приходится использовать иные лексемы: *ничто произошло* (им. п.), *увидел ничто* (вин. п.), *но приблизился к чему-то* (дат. п.), *подумал о чём-то* (предл. п.).

Кроме того, последовательное «открытие» падежей для местоименных существительных *он, она, оно* должно привести к увеличению числа паде-

жей. Местоименные существительные *он, она, оно* в ряде падежей, выделенных для существительных, могут иметь по две формы, отнюдь не находящиеся между собой в отношениях свободного варьирования.

Речь идет о случаях типа *у него нет* (род. п.) и *лишиться его* (род. п.), *подошёл к ней* (дат. п.) и *нравится ей* (дат. п.), *вижу их* (вин. п.) и *смотрю на них* (вин. п.) и т. д. Несколько упрощая дело, можно утверждать, что после предлогов выступают формы с начальным *н*, при отсутствии предлогов — формы, не имеющие *н*.

Наличие разных форм характеризует не только местоименные существительные *он, она, оно, они*, но и *никто, ничто, некого, нечего*, у которых в сочетании с предлогом представлены так называемые разорванные формы: *не лишился ничего (никого)*; *нет ни у чего (ни у кого)*; *сказать некому — пойти не к кому*.

#### 5.6.4. Местоименные прилагательные

Значительную группу лексем среди традиционно выделяемых местоимений составляют местоименные прилагательные. По своей семантике они подразделяются на притяжательные: *мой, твой, наш, ваш, свой*; указательные: *тот, этот, такой, этаким, таков, следующий*; определительные: *всякий, всяческий, каждый, любой, весь, целый, иной, другой, сам, самый*; вопросительные: *какой, который, чей, каков*; неопределенные: *какой-то, какой-нибудь, какой-либо, чей-то, чей-нибудь, чей-либо, кое-какой, некоторый, некий*; отрицательные: *никакой, никоторый, ничей*.

Все местоименные прилагательные — изменяемые, они обладают словоизменительными категориями рода, числа и падежа. Местоименные прилагательные лишены противопоставления по полноте—краткости. Все они полные в том смысле, что изменяются не только по числам и родам, но и по падежам. В отличие от обычных прилагательных не все местоименные прилагательные могут выступать в роли сказуемого.

#### 5.6.5. Местоименные числительные

Местоименные слова выделяются и среди числительных: *сколько, столько, несколько, сколько-нибудь, сколько-то, нисколько*. Так же, как и иные числительные, эти слова не обладают категориями рода и числа.

Проявление категорий падежа у этих лексем такое же, как у обычных числительных: они «подчиняют» себе существительные в именительном падеже (*несколько друзей*), равном ему винительном (*вижу несколько до-*

мов) и в дательном падеже с предлогом *по* в распределительном значении: *По сколько рублей они получили?*

Во всех остальных падежах местоименные числительные согласуются в падеже со связанным с ними существительным: *к нескольким представителям, говорить о стольких вещах*. У местоименных числительных в винительном падеже может проявляться категория одушевленности — неодушевленности: *(вижу) несколько домов — нескольких юношей*. Впрочем, это противопоставление не всегда выдерживается: можно сказать и *вижу несколько юношей*. По семантике местоименные числительные могут быть подразделены на вопросительные: *сколько*; указательные: *столько*; отрицательные: *нисколько*; неопределенные: *несколько, сколько-то* и т. п.

#### 5.6.6. Склонение местоименных слов

Достаточно полное описание типов склонения местоименных слов обычно представлено в грамматических словарях русского языка.

Нерегулярное склонение имеют местоименные существительные *я, ты, мы, вы, себя*:

И.	<i>я</i>	<i>ты</i>	<i>мы</i>	<i>вы</i>	
Р., В.	<i>меня</i>	<i>тебя</i>	<i>нас</i>	<i>вас</i>	<i>себя</i>
Д.	<i>мне</i>	<i>тебе</i>	<i>нам</i>	<i>вам</i>	<i>себе</i>
Т.	<i>мной</i>	<i>тобой</i>	<i>нами</i>	<i>вами</i>	<i>собой</i>
П.	<i>мне</i>	<i>тебе</i>	<i>нас</i>	<i>вас</i>	<i>себе</i>

Склонение местоименных существительных *кто, что* приближается к склонению прилагательных типа *лисье* или *дядино*. Местоименные существительные *он, она, оно* имеют различные основы для именительного падежа обоих чисел и для всех иных падежей. Окончания таких лексем практически совпадают с окончаниями прилагательных типа *лисий*. Особенность данных лексем состоит в том, что в зависимости от наличия или отсутствия предлогов они имеют разные падежные формы.

Склонение местоименных существительных *никто, ничто*, а также *некого, нечего* (обладающих дефектом в парадигме) управляется двумя правилами: первое относится к склонению местоименных существительных *кто, что*, второе касается различий в формах одного падежа в зависимости от наличия или отсутствия предлога.

И.		<i>он, оно</i>	<i>она</i>	<i>они (мн. ч.)</i>
Р.	без предл.	<i>его</i>	<i>её</i>	<i>их</i>
	с предл.	<i>от него</i>	<i>от неё</i>	<i>от них</i>
Д.	без предл.	<i>ему</i>	<i>ей</i>	<i>им</i>
	с предл.	<i>к нему</i>	<i>к ней</i>	<i>к ним</i>
В.	без предл.	<i>его</i>	<i>её</i>	<i>их</i>
	с предл.	<i>про него</i>	<i>про неё</i>	<i>про них</i>
Т.	без предл.	<i>им</i>	<i>ею (ей)</i>	<i>ими</i>
	с предл.	<i>с ним</i>	<i>с ней (нею)</i>	<i>с ними</i>
П.		<i>о нем</i>	<i>о ней</i>	<i>о них</i>

Среди местоименных прилагательных выделяется несколько разновидностей склонения, напоминающих склонение прилагательных типа *дядин* или *лисий*.

## 6. МИНИМАЛЬНЫЕ СХЕМЫ ПРЕДЛОЖЕНИЙ

Расширенные схемы предложения = минимальные схемы + не входящие в них конструктивные, то есть существенные для семантической структуры предложения, компоненты.

Для описания объективного содержания предложения в современных исследованиях используются разные понятия, например, понятие пропозиции, заимствованное из логики. Структуру пропозиции определяет предикат, который несет в себе признак предмета или отношение между предметами, указывает определенные места для предметов — актантов, определяет их количество и роли.

Предикаты могут быть нульместными, одноместными и многоместными. Их валентные свойства связаны с семантическими и определяются ими. Так, предикаты **состояния и характеристики** лица, животного или предмета — одноместны. Их единственный актант всегда имеет значение **субъекта состояния**: *Он тревожится (тревожен, в тревоге); Ему холодно.*

Предикаты **состояния окружающей среды**, нульместны: *Светает; Заря; Морозно; Идет дождь.*

Предикаты **активного действия** могут обозначать **процессы**, то есть действия, не направленные на изменение ситуации, в которой они протекают: *Дети бегают; Все работают* — в этом случае, они одноместны; а мо-

гут обозначать **события**, то есть действия, направленные на изменение ситуации, предполагающие «исходное» и «завершающее» ее состояния: *Художник нарисовал картину; Мы сажали капусту; Я взял у соседа книгу* — в этом случае они многоместны.

Количество мест также зависит от семантики предиката, равно как и от роли актантов. Так, предикаты со значением **давания—получения** трехместные: *Сосед дал мне книгу*.

Ниже перечисляются минимальные, т.е. простейшие, схемы предложений.

### **Двухкомпонентные номинативные**

1. Существительное (в ИП) + Глагол (его спрягаемая форма).  
 $N_1 V_f$ . Например: *Грачи прилетели; Зеленеют деревья; Все дела делаются людьми*.

С помощью предикативной конструкции этой схемы можно выразить пропозиции с одноместным или двухместным предикатом, имеющим значение активного целенаправленного действия (то есть здесь можно рассматривать одноместные или двухместные предикаты указанного выше типа): *Здесь он работает; Утром он разносит газеты*.

По этой схеме могут быть оформлены пропозиции с одноместным предикатом характеризующей семантики: *Он лгун; Он добряк*.

2. Существительное (в ИП) + Глагол-связка «быть» (его спрягаемая форма) + Прилагательное или Страдательное Причастие (их полная форма в ИП или ТП, или краткая форма или компаратив).  
 $N_1 \text{Cop}_f \text{Adj}_f /_{1/5}$ . Например: *Ночь была тиха (тихая, тихой); Через час был объявлен привал; Машины готовы к испытаниям; Он ранен*.

По этой схеме могут быть оформлены пропозиции с одноместным предикатом характеризующей семантики: *Он лжсет; Он хромает*.

3. Существительное (в ИП) + Глагол-связка «быть» (его спрягаемая форма) + Существительное (в ИП или ТП).  
 $N_1 \text{Cop}_f N_{1/5}$ . Например: *Он был студент (студентом); Орел — хищник; Это наше общежитие*.

По этой схеме могут быть оформлены пропозиции с одноместным предикатом характеризующей семантики: *Он лживый; Он хромой; Он добрый*.

4. Существительное (в ИП) + Глагол-связка «быть» (его спрягаемая форма) + Существительное (в беспредложной или предложной

форме любого косвенного падежа (все падежи, кроме ИП и ТП), которая способна сочетаться с глаголом-связкой «быть») или Наречие (способное сочетаться с глаголом-связкой «быть»).

$N_1 \text{Cor}_f N_2 \dots_{\text{pr}} / \text{Adv}_{\text{pr}}$ . Например: *Этот дом будет без лифта; Мы были в отчаянии; Чай — с сахаром; Приход Ивана Ивановича был кстати; Все были на чеку; У него глаза навывкате.*

### Двухкомпонентные инфинитивные

5. Глагол (инфинитив) + Глагол (его спрягаемая форма).  
 $\text{InfV}_f$ . Например: *Не мешало б нам встречаться чаще (Светлов); Отмалчиваться не следует; Курить воспрещалось; Быть космонавтом (смелым) хочется каждому мальчишке; Друзьям разрешалось быть вместе.*
6. Глагол (инфинитив) + Глагол-связка «быть» (его спрягаемая форма) + Прилагательное или Страдательное Причастие (их полная форма в ИП или ТП, или краткая форма или компаратив).  
 $\text{InfCor}_f \text{Adj}_f /_{1/5}$ . Например: *Промолчать было разумно (разумнее, самое разумное, самым разумным); Уговаривать его было излишне (излишнее, излишним); Нужно уехать; Правильнее было бы признать свою ошибку.*
7. Глагол (инфинитив) + Глагол-связка «быть» (его спрягаемая форма) + Существительное (в ИП или ТП).  
 $\text{InfCor}_f N_{1/5}$ . Например: *Дозвониться — проблема (было проблемой); Главной его целью было (главная его цель была) увидеть все своими глазами; Строить — это радость; Любить иных — тяжёлый крест (Пастернак); Оказывается, быть взрослой — не всегда преимущество (Нагибин); Превосходная должность быть на земле человеком (Горький).*
8. Глагол (инфинитив) + Глагол-связка «быть» (его спрягаемая форма) + Существительное (в беспредложной или предложной форме любого косвенного падежа, которая способна сочетаться с глаголом-связкой «быть») или Наречие (способное сочетаться с глаголом-связкой «быть»)).  
 $\text{InfCor}_f N_2 \dots_{\text{pr}} / \text{Adv}_{\text{pr}}$ . Например: *Промолчать было не в его правилах; Купить машину нам не по средствам; Молчать некстати; Идти дальше было невозможно.*

9. Глагол (инфинитив) + Глагол-связка «быть» (его спрягаемая форма) + Глагол (инфинитив).

InfCop<sub>г</sub>Inf. Например: *Отказаться было обидеть; Быть студентом — это постоянно учиться мыслить; Быть актером — прежде всего быть талантливым человеком.*

### Однокомпонентные

10. Глагол (его форма единственного числа третьего лица или среднего рода).

V<sub>s3/n</sub>. Например: *Скрипело, свистало и выло в лесу (Заболоцкий); Смеркается; Ему нездоровится; Дохнуло свежестью; Крышу охватило пламенем; Пароход покачивало; У него накопело на сердце; Об этом уже писалось.*

Схема предполагает изображение семантического признака нецеленаправленности, произвольности действия, несвязанности его с сознательной деятельностью человека.

11. Глагол (его форма множественного числа третьего лица).

V<sub>pl3</sub>. Например: *За столом зашумели; Его обидели; Здесь о молодых специалистах заботятся, им доверяют; Во время еды не разговаривают.*

Эта схема изображает действие как сознательное, целенаправленное и потому исходящее обязательно от человека, но лицо, от которого исходит это действие, остается «за кадром». Поэтому она не допускает употребления глаголов, обозначающих действия и состояния, не связываемые с человеком (*годиться, зазеленеть, случаться, отразиться, состоять из*).

Здесь можно рассматривать одноместные или двухместные предикаты, имеющие значение активного целенаправленного действия: *Здесь работают; Утром разносят газеты.*

12. Глагол-связка «быть» (его форма единственного числа третьего лица или среднего рода) + Прилагательное или Страдательное Причастие (их краткая форма или компаратив единственного числа среднего рода).

Cop<sub>s3/n</sub>Adj<sub>fsn</sub>. Например: *Было темно; Морозно; Ночью будет холодно; Душно без счастья и воли (Некрасов).*

По этой схеме строятся предложения, передающие разнообразное состояние.

13. Глагол-связка «быть» (его форма единственного числа третьего лица или среднего рода) + Существительное (в беспредложной или предложной форме любого косвенного падежа, которая способна сочетаться с глаголом-связкой «быть») или Наречие (способное сочетаться с глаголом-связкой «быть»).

$\text{Cop}_{s3/n} \text{N}_2 \dots_{\text{pr}} / \text{Adv}_{\text{pr}}$ . Например: *Было уже полночь; Завтра будет без осадков; Нам не до сна; Ей было невдомек; Пусть будет по-твоему; Ему не к спеху.*

14. Глагол-связка «быть» (его форма множественного числа третьего лица) + Прилагательное или Страдательное Причастие (их краткая форма или компаратив множественного числа).

$\text{Cop}_{\text{pl3}} \text{Adj}_{\text{gr1}}$ . Например: *Ему были рады; Им довольны; Отказом были обижены.*

15. Глагол-связка «быть» (его форма множественного числа) + Существительное (в беспредложной или предложной форме любого косвенного падежа, которая способна сочетаться с глаголом-связкой «быть») или Наречие (способное сочетаться с глаголом-связкой «быть»).

$\text{Cop}_{\text{pl}} \text{N}_2 \dots_{\text{pr}} / \text{Adv}_{\text{pr}}$ . Например: *Дома были в слезах; От него были в восторге; С ним были запросто.*

16. Глагол-связка «быть» (его спрягаемая форма) + Существительное (в ИП).

$\text{Cop}_{\text{r}} \text{N}_1$ . Например: *Шепот. Робкое дыхание. Трели соловья (Фет); Тишина; Была зима.*

Здесь можно рассматривать одноместные или двухместные предикаты, имеющие значение активного целенаправленного действия: *Здесь ему работать; Утром ему разносить газеты.*

17. Глагол (инфинитив).

$\text{Inf}$ . Например: *Не нагнать тебе бешеной тройки (Некрасов); Только детские книги читать. Только детские думы лелеять (Мандельштам); Быть рекам чистыми; Быть мальчишке поэтом; Быть по-вашему; Всем быть в спортивной форме.*

Здесь можно рассматривать одноместные или двухместные предикаты, имеющие значение активного целенаправленного действия: *Здесь у него работа; Утром у него разноска газет.*

**СПИСОК ЛИТЕРАТУРЫ**

1. **Апресян Ю.Д.** Экспериментальное исследование семантики русского глагола. — М.: Наука, 1967. — 251 с.
2. **Мельчук И.А.** Опыт теории лингвистических моделей типа «Смысл $\leftrightarrow$ Текст». — М. 1974. — 315 с.
3. **Маркус С.** Теоретико-множественные модели языков. — М.: Наука, 1970. — 332 с.
4. **Современный русский язык:** Учеб. для филол. спец. высших учебных заведений/ Под ред. В.А. Белошапковой. — М.: Азбуковник, 1997. — 928 с.
5. **Современный русский язык:** Учеб. для филол. спец. высших учебных заведений/ Под ред. Д.Э. Розенталя. — М.: Изд. МГУ, 1971. — 636 с.