

ФОРМАЛНА ГРАМАТИКА НА БЪЛГАРСКИЯ ЕЗИК

**Петя Осенова
Кирил Симов**

**Институт за паралелна обработка на информацията, БАН
София, 2007**



Тази книга се издава благодарение на финансирането, получено по два проекта на Института за паралелна обработка на информацията (ИПОИ) на БАН. Първият проект BulTreeBank, с ръководител Кирил Симов, беше реализиран от Секцията за лингвистично моделиране в периода 2001 – 2004 г. Този проект осигури възможността да се създаде интердисциплинарен колектив, който да работи по създаването на базата със синтактични описания и по разработването на компютърната среда, осигурила унифицирано аотиране на 15 000 изречения на български език. Вторият проект, БИС-21++ – така нареченият Център по компетентност в ИПОИ – подпомага обучението на млади учени и издаването на учебни материали. Благодарение на подкрепата на БИС-21++ настоящата книга се публикува за първи път.

Д-р Петя Осенова е главен асистент в Катедрата по български език, СУ „Св. Климент Охридски“. Тя е дългогодишен сътрудник на ИПОИ – БАН и работи по проекти на Секцията за лингвистично моделиране. Нейните главни научни интереси са в областта на формалната и корпусната лингвистика. Тя има редица публикации (самостоятелно и в съавторство) на международни форуми и в научни списания, които дискутират лингвистични проблеми с оглед на компютърната обработка на българския език.

Д-р Кирил Симов е старши научен сътрудник II ст. в Секцията за лингвистично моделиране на ИПОИ – БАН. През последните 20 години той се занимава с модели на българския език, построени чрез логически формализми за представяне на лингвистични знания. Ръководил е няколко съвместни проекта с Университета в Тюбинген, финансирани от фондация Фолксваген (Германия), както и редица научноизследователски проекти по 6-тата и 7-мата рамкова програма на Европейската комисия. Кирил Симов е член на Борда на Международното общество за Опорната фразова граматика (от 2002 г. досега) и на Борда на Европейската асоциация по компютърна лингвистика (за 2007 и 2008 г.).

ФОРМАЛНА ГРАМАТИКА НА БЪЛГАРСКИЯ ЕЗИК

**Петя Осенова
Кирил Симов**

**първо издание
ISBN 978-954-92148-2-6**



Съдържание

1	Увод.....	5
2	Опорна фразова граматика – формални основи и универсални принципи	7
2.1	Йерархия на типовете	10
2.2	Граматични принципи	13
2.3	Опорната фразова граматика – йерархия и принципи	13
3	Представяне на формалния модел на ОФГ в графичен вид	22
4	Лексикални елементи	30
4.1	Лексикален елемент глагол	31
4.2	Лексикален елемент причастие	39
4.3	Лексикален елемент деепричастие	40
4.4	Лексикален елемент съществително име	41
4.5	Лексикален елемент местоимение	41
4.6	Лексикален елемент прилагателно	42
4.7	Лексикален елемент числително	43
4.8	Лексикален елемент фамилно име или прилагателно, образувано от собствени имена	44
4.9	Лексикален елемент наречие	44
4.10	Лексикален елемент съюз	45
4.11	Лексикален елемент предлог	46
4.12	Лексикален елемент частица	47
4.13	Лексикален елемент междуметие	47
4.14	Промяна на типа (type shifting)	47
5	Глаголна фраза (вербална фраза)	50
5.1	Глаголна фраза от типа опора-комплемент – VPC	51
5.2	Глаголна фраза от типа опора-подлог – VPS	57
5.3	Глаголна фраза от типа опора-адюнкт – VPA	59
5.4	Глаголна фраза от типа опора-запълваща част – VPF	62
5.5	Подчинени изречения	63

6 Именна фраза (номинална фраза)	71
6.1 Именна фраза от типа опора-комплемент – NPC	73
6.2 Именна фраза от типа опора-адюнкт – NPA	74
6.3 Типове адюнкти в именните фрази	76
7 Фраза на прилагателното (адекативна фраза)	81
7.1 Фраза на прилагателното от типа опора-комплемент – APC	81
7.2 Фраза на прилагателното от типа опора-адюнкт – APA	82
8 Фраза на наречието (адвербиална фраза)	84
8.1 Фраза на наречието от типа опора-комплемент – AdvPC	84
8.2 Фраза на наречието от типа опора-адюнкт – AdvPA	85
9 Предложна фраза	86
9.1 Предложна фраза от типа опора-комплемент – PP(C)	86
9.2 Предложна фраза от типа опора-адюнкт – PP(A)	89
10 Координационни фрази	91
11 Елиптически конструкции	96
11.1 Елипси с елипсни носители в рамките на изречението	97
11.2 Елипси с елипсни носители извън изречението	101
12 Прагматични фрази	103
13 Кореференция	111
14 Заключение	113
Използвана литература	114
Приложение 1	116
Приложение 2	118

Нашата книга има за основа направеното по проекта BulTreeBank (2001-2004)¹. Това е съвместен изследователски проект между Института по паралелна обработка на информацията, БАН и Университета в Тюбинген, Германия.

Този проект създаде представителна база от формални синтактични анализи на български изречения. Тя е в синхрон със създадените подобни бази за други езици: английски, немски, чешки, турски, френски, испански и т.н.

Нашата задача по тази инициатива беше да изградим граматичен модел на основата на голям корпус от примери както от художествената литература, така и от периодичните издания. В този смисъл граматиката е изградена върху конкретни примери от реални текстове. Предимство е, че се разглеждат явления, типични не само за изолирани изречения, но и за свързан текст. Освен това използвахме вече описаните примери в трети том на Граматика на съвременния български книжовен език и в трудовете по синтаксис на Йордан Пенчев и Стефан Брезински, за да създадем основно множество от изречения, което да обхваща регистрираните синтактични явления.

Създаването на компютърен модел на граматиката поставя изискването за завършеност (цялостност) на модела, т.е. всяко явление или фраза трябва да получи адекватно и непротиворечиво представяне в общ формален модел. По този начин се избягва „премълчаването“ или случайното пропускане на трудни за решаване случаи. Построяването обаче на такъв цялостен модел на граматиката все още не е постижимо, тъй като езикът е сложна система, върху която влияят различни фактори от лингвистичен и екстралингвистичен характер. По тази причина на някои места нашите решения не са окончателни, а са представени като една от възможностите при текущото състояние на изградения от нас формален модел на граматиката на българския език. В тези случаи ние представяме и някои алтернативни решения, като често дискутираме техните предимства и/или недостатъци.

Разбира се, ние не започваме формалното моделиране на българския език от нулата. Усилия в тази насока е имало от 60-те години насам. Без претенция за изчерпателност ще споменем имената на учени като Александър Людсканов, Мирослав Янакиев, Елена Паскалева, Корнелия Илиева, които са описали различни явления с помощта на алгоритмични средства. Също така не можем да пропуснем трудовете по синтаксис на Йордан Пенчев, който от 1984 г. нататък слага началото на първите описания на българския език в духа на теориите на Чомски. Автор, който поставя синтаксиса във връзка с други нива на езика (което е основна идея в използваната от нас теория), е Руселина Ницолова. Тя например въвежда моделирането на синтаксиса в прагматичен план с труда си „Прагматичен аспект на изречението в българския книжовен език“ от 1984 г.

За формализиране на данните използваме апарата на т.нар. Опорна фразова граматика², която представлява лексикална граматична теория с формално представяне.

¹ Този проект е по програмата „Cooperation with Natural and Engineering Scientists in Central and Eastern Europe“ на фондация „Фолксваген“, Германия, договор номер I/76 887.

² Оригинално заглавие на английски е: Head-driven Phrase Structure Grammar. Основополагаща роля за създаването на тази лингвистична теория имат Ivan Sag и Carl Pollard.

Тъй като формалният апарат е сложен, ние предлагаме опростена версия с оглед на целите на изложението ни. Важно е да подчертаем, че въпреки опростяването предлаганите решения запазват своята валидност. Използвайки формален апарат, ние постигаме основната цел на една цялостна граматика – да опише и обобщи лингвистичните факти по адекватен и непротиворечив начин.

Книгата предлага богат илюстративен материал, което, надяваме се, ще улесни възприемането на описанията ни от читателите. Тя е предназначена за всички, които се интересуват от принципите на моделиране на езика и които не се страхуват да търсят нови предизвикателства в езиковите интерпретации.

Разделът **Именна фраза (номинална фраза)** е авторска разработка на Петя Осенова. Разделът **Опорна фразова граматика – формални основи и универсални принципи** е авторска разработка на Кирил Симов. Останалите раздели са разработени съвместно от двамата автори.

* * *

Благодарим на партньорите си от Катедрата по компютърна лингвистика към Тюбингенския университет с ръководител проф. Ерхард Хинрикс за ползотворното сътрудничество по проекта. Задължени сме на немската фондация „Фолксваген“, без чиято финансова подкрепа този проект нямаше да бъде осъществен.

Благодарим на всички участници в проекта, които допринесоха с труда и ентузиазма си той да стане реалност. Това са: лингвисти – доц. Сия Колковска, д-р Елисавета Балабанова, Магделена Паунова, Димитър Дойков, д-р Илияна Гаравалова, Милена Славчева; информатици – Александър Симов, д-р Милен Куйлеков, Красимира Иванова.

Благодарим на колегите си в чужбина, с които сме дискутирали адекватността на анализите за специфични езикови явления. Сред тях са: Франк Рихтер, Манфред Зайлър, Ерхард Хинрикс, Беата Травински, Адам Пржепиорковски.

Бихме искали да изкажем благодарност и на колегите си в София, които споделиха мненията и препоръките си по английската версия на този труд през януари 2005 г. на дискусия в Софийския университет „Св. Климент Охридски“. Това са: проф. Йордан Пенчев, проф. Руселина Ницолова, ст.н.с. II ст. Елена Паскалева, доц. Радка Влахова, доц. Йовка Тишева.

Отделно бихме искали да благодарим на ст.н.с. II ст. Галя Ангелова, проф. Руселина Ницолова, доц. Йовка Тишева и Ласка Ласкова за изключително полезните забележки, коментари и предложения по ранните редакции на части от тази книга.

Разбира се, всички евентуални грешки и неточности в текста остават единствено наша отговорност!

Изказваме специалната си благодарност към проекта BIS21++, финансиран от Европейската комисия чрез договор INCO-CT-2005-016639, с чието любезно съдействие стана възможно издаването на тази книга.

септември 2007

Авторите

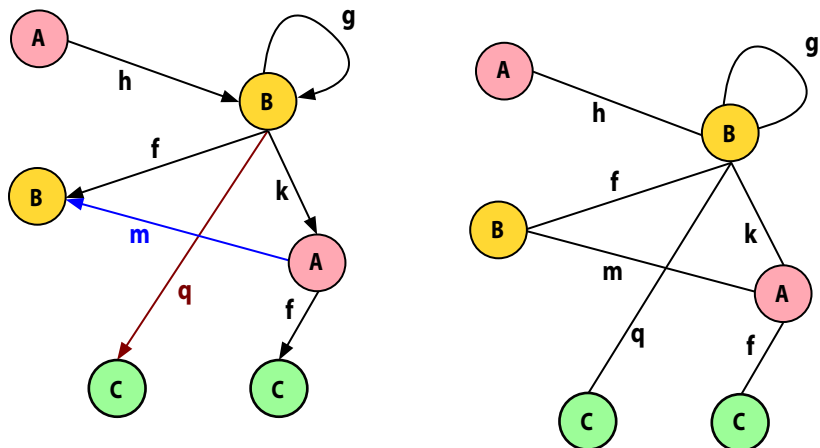
Опорна фразова граматика – формални основи и универсални принципи

В тази глава представяме накратко основните характеристики и принципи на лингвистичната теория **Опорна фразова граматика** (ОФГ) (Head-driven Phrase Structure Grammar – вж. (Pollard and Sag 1994)).

В основата на ОФГ стои разбирането, че езикът се състои от **езикови обекти**, които оформят езиковата действителност. **Граматиката** на даден език е теория за езиковата действителност. В качеството си на теория граматиката разпознава кои обекти са от езиковата действителност и каква е тяхната връзка с други обекти от тази действителност. **Лингвистичната теория** (в нашия случай ОФГ) дефинира вида на граматиките и начина на свързване на елементите на граматиката с езиковата действителност. ОФГ прави абстракция над езиковата действителност, разглеждайки езика като система от типове, които отразяват свойствата и взаимовръзките между езиковите обекти. Тези типове се представят чрез **атрибутни структури с функционални атрибути**. Атрибутните структури с функционални атрибути са специален вид **графи** (вж. по-нататък за дефиниция на граф). Една ОФГ граматика се състои от **лингвистична онтология**, представена като **йерархия на типове**, и от **граматични принципи**, представени като **граматична теория**. Йерархията на типовете определя видовете езикови обекти и техните основни свойства. В духа на атрибутните структури йерархията определя възможните типове за възлите на тези структури и типовете на дъгите, които излизат от тях. **Граматичната теория** е множество от ограничения, дефинирани над йерархията на типовете. Тези ограничения определят структурата на езиковите обекти в езика. Това става, като от обектите, които се допускат от йерархията на типовете, се избират онези, които удовлетворяват ограниченията на теорията. Един от основните механизми за определяне на структурата на лингвистичните обекти е **структурното съвместяване** (*reentrance* или *structure sharing*). Най-общо, структурното съвместяване означава, че даден езиков обект е свързан с друг езиков обект чрез две различни характеристики, които се представят чрез последователност от атрибути. ОФГ се нарича лексикална лингвистична теория, тъй като основните ограничения се задават към думите на дадения език в частта на граматичната теория, която се нарича **лексикон**.

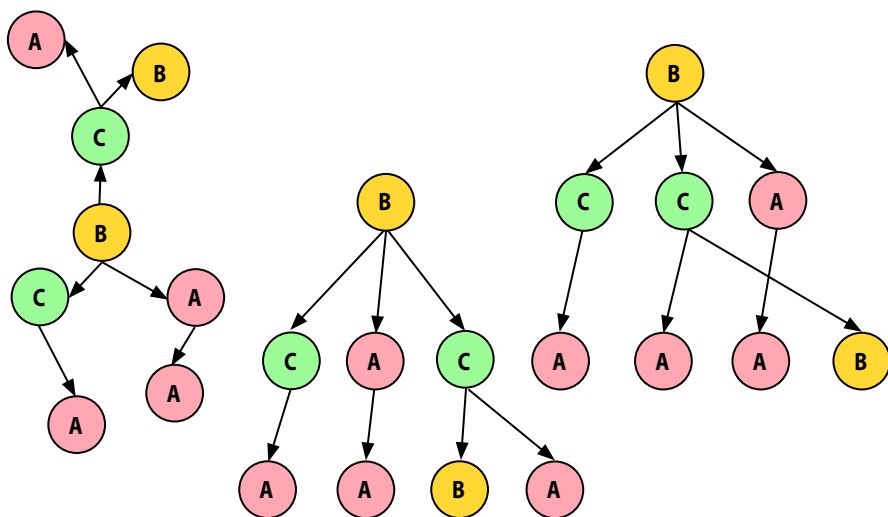
Следват основните понятия от теория на графите, които ще използваме в книгата. Граф G е наредена двойка $\langle V, A \rangle$, където $V = \{n_1, n_2, n_3, \dots, n_k\}$ е множество от **възли (върхове)** на графа G , A е множество от двойки елементи на V , наречени **дъги (ребра)** на графа G . Ако дъгите са наредени двойки, тоест знае се, кой елемент е първи и кой втори, то графът се нарича **насочен (ориентиран)**, ако дъгите не са наредени двойки, то графът е **ненасочен (неориентиран)**.

При неориентираните графи приемаме, че всяка дъга е представена в двете възможни подредби на краищата си – $\langle n_i, n_j \rangle$ и $\langle n_j, n_i \rangle$. Път в графа G е редица от дъги от вида $\langle n_{i_1}, n_{i_2} \rangle, \langle n_{i_2}, n_{i_3} \rangle, \dots, \langle n_{i_j}, n_{i_k} \rangle$. Първият възел на първата дъга в пътя е неговото **начало**, а вторият възел на последната дъга в пътя е неговият **край**. **Дължина** на пътя е броят на дъгите в него. **Цикъл** е път със съвпадащи начало и край. Един граф е **цикличен**, ако съдържа поне един цикъл. В противен случай графът е **ацикличен**. Допълнително към възлите и дъгите могат да бъдат прикрепяни етикети, които ги наименоуват. Графът G е **краен**, ако множествата на възлите и дъгите са крайни. Графът G е **свързан**, когато между всеки два негови възела съществува път. Визуално графите се представят като множество от точки в равнината (толкова, колкото са възлите в графа), съпоставени на възлите от графа, и линии, свързващи две точки, ако съответстващата им дъга е от графа. Линиите завършват със стрелка, ако графът е насочен. Когато възлите са наименоувани с етикети, възлите се представят като овали и етикетите се записват вътре. Когато дъгите са наименоувани, етикетите се изписват в близост до съответните дъги. При графичното представяне ние също така използваме оцветяване на овалите и дъгите с цел по-добро ориентиране в графите, но самите цветове нямат отношение към формалните характеристики на съответните графи. Следват два примера на един насочен и един ненасочен граф. И двата графа са циклични, с наименоувани дъги и възли:

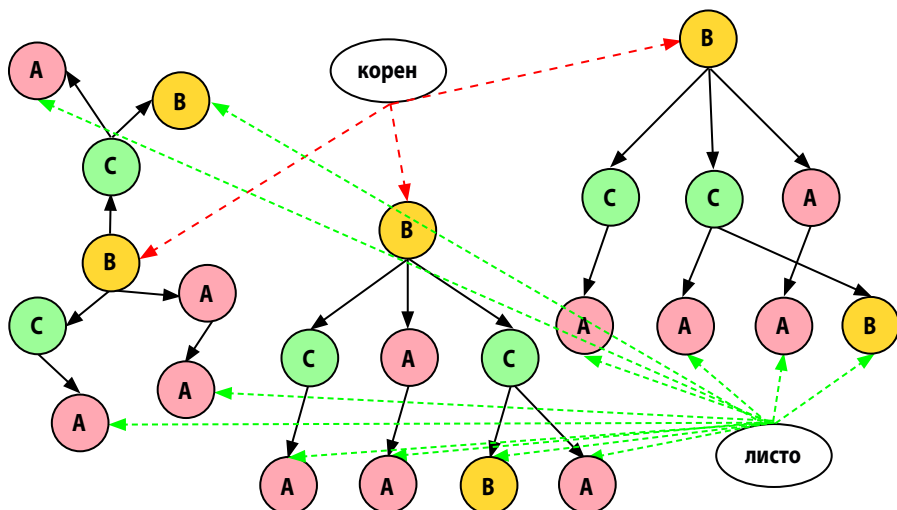


Дърво е свързан граф без цикли. Съществуват насочени и ненасочени дървета. Тук ще работим само с насочени дървета. За дадена дъга $\langle a, b \rangle$ в дадено дърво, a се нарича **родител** на b и b е **дете** (пряк наследник) на a . Децата на децата (и така нататък) на даден възел a се наричат **непреки наследници** на a . Множеството на преките и непреки наследници на даден възел се наричат просто **наследници**. **Корен** на дадено дърво е единственият възел, който не е дете на друг възел. **Листа** на дървото са възлите, които нямат наследници. В нашата работа ще използваме специален вид дървета, които ще наричаме **наредени дървета**. При тях за всеки възел съществува наредба на неговите

деца. Дори ще наложим допълнително условие да съществува наредба на листата на дървото. Следват няколко примера на дървета.



Първото дърво е представено в свободна наредба на възлите. Второто и третото дърво са наредени – при второто наследниците на корена са в последователност $\langle C, A, C \rangle$, а наследниците на второто C са в последователност $\langle B, A \rangle$; при третото наследниците на корена са в последователност $\langle C, C, A \rangle$, а листата са в последователност $\langle A, A, A, B \rangle$, която налага две от дъгите да се пресичат. Така третото дърво показва и наредба на листата си. Второто и третото дърво също така показват вертикално представяне от корена към листата. В следващата схема посочваме още кои възли са корените и листата на горните три дървета.



Както беше казано по-горе, ОФГ използва специален вид графи, които са насочени, възлите и дъгите са наименувани по такъв начин, че от един възел излиза точно една дъга с даден етикет. ОФГ разглежда този специален вид графи като абстракция над езиковите обекти в езиковата действителност. В следващите раздели се описват механизмите, които ОФГ използва, за да дефинира тези графи. Първата стъпка е да се дефинират етикетите, които да се използват за възлите и дъгите. Втората стъпка е да се дефинира структурата на графите. В текста, който следва, ще говорим за езикови обекти, но ще използваме съответните графи като техен модел.

2.1

Йерархия на типовете

Йерархията на типовете определя потенциалните видове езикови обекти. Йерархията дефинира самите **типове**, **релация на наследяване** и **функция на съвместимост**, определяща **атрибутите**, които са съвместими с (подходящи за) даден тип. Интерпретацията на йерархията на типовете е следната. Първо, тя е свързана с област от обекти (в нашия случай, това са езиковите обекти). Всеки тип се интерпретира като множество от езикови обекти от областта на интерпретацията. За всеки един такъв обект ще казваме, че е от дадения тип. Всеки атрибут, който се отнася за даден тип T , обозначава връзка (функция), дефинирана за езиковите обекти от типа T . Тази връзка ги съединява с обекти, които са подходящи за този атрибут на тип T . Наследяването е монотонно, т.е. един атрибут, дефиниран за даден тип, се наследява от всичките му подтипове и неговата стойност се ограничава от същия или по-специфичен тип при наследяване. Наследяването е многомерно, тоест даден тип може да наследява от повече типове. Тук представяме йерархията на типовете по следния начин: първо представяме съответния тип (използваме малки латински букви в курсив – *type*), на следващите редове с определено отместване се декларират атрибутите, които са подходящи за този тип (използваме главни латински букви в полчер за атрибутите – **FEAT**). Всяка декларация на атрибут има следния формат: **FEAT: type**, където **FEAT** (от feature) е атрибут и *type* е тип. Всяка декларация на атрибут означава, че даденият атрибут е подходящ за дадения тип и неговата стойност трябва да е от типа, представен в декларацията. Ако даден тип няма подходящи атрибути, то към него не се задават декларации на атрибути. Подтиповете на даден тип се задават след декларациите на атрибути (ако има такива) с определено отместване надясно. Подйерархията за даден подтип се задава след него. Всеки подтип наследява атрибутите на неговите надтипове, но може да има и собствени атрибути, които се декларират локално. Когато при наследяване даден атрибут има по-специфична стойност, тя се декларира също локално. Ето един пример:

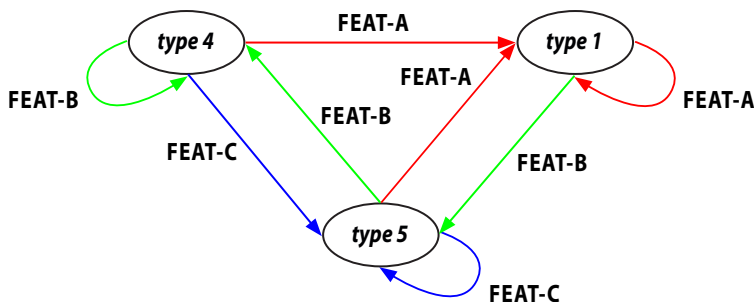
```

root
  FEAT-A : type1
  FEAT-B : type2
    type1
    type2
      FEAT-B : type4
      FEAT-C : type2
        type4
        type5
      type3

```

Тук има шест типа (*root*, *type1*, *type2*, *type3*, *type4*, *type5*) и три атрибута (**FEAT-A**, **FEAT-B**, **FEAT-C**). Типът *root* има два подходящи атрибута: **FEAT-A** и **FEAT-B**. Те имат като ограничения върху стойностите си съответно типовете *type1* и *type2*. Типът *root* има три подтипа: *type1*, *type2* и *type3*. Всеки от тях наследява двата атрибута, декларирани за типа *root*. Типът *type2* има допълнителен подходящ атрибут **FEAT-C** и допълнително ограничение на стойността на атрибута **FEAT-B**, което е зададено с типа *type4*. Типът *type2* има два подтипа: *type4* и *type5*. Те наследяват от типа *type2* всичките три атрибута: **FEAT-A**, **FEAT-B** и **FEAT-C**. Наследяването е транзитивна релация³. Типовете *type1*, *type3*, *type4* и *type5* са минимални по отношение на обектите, които обозначават, или, с други думи, са максимално специфични по отношение на информацията, която представят.

Йерархията на типовете играе много важна роля при дефинирането на езиковите обекти. Всеки обект принадлежи на точно един минимален тип. Ако даден атрибут е дефиниран като подходящ за даден тип, то всеки обект от този тип притежава този атрибут и неговата стойност (съответният обект) е от типа, който е ограничение на стойността на атрибута в съответната декларация на атрибути. В това отношение йерархията на типовете предопределя възможните езикови обекти. Действителните езикови обекти зависят допълнително и от принципите на граматиката, които са описани в следващите раздели. За да покажем връзката на обектите, допустими спрямо тази йерархия на типовете и атрибутните структури, които представляват техен модел, ще покажем граф, който е построен на базата на йерархията:



³ R е транзитивна, ако от $R(x,y)$ и $R(y,z)$ следва $R(x,z)$.

В този граф имаме три възела (съответстващи на три обекта) – по един от тип *type1*, *type4* и *type5* (всеки един от тях е минимален). От съответните възли излизат съответен брой дъги в зависимост от декларациите в йерархията. За тип *type1* – две дъги и за другите два типа по три дъги. Дъгите свързват съответните възли с възли от подходящ тип. Например, дъгата с етикет **FEAT-B**, излизаща от възела с етикет *type1*, го свързва с възел от тип *type5*, тъй като *type5* е подходящ за този атрибут за типа *type1*. Тук трябва да се има предвид наследяването за *type1* на атрибута **FEAT-B** и фактът, че *type5* е подтип на *type2*. Използваме цветове за по-ясно открояване на съответните дъги. Друг начин за представяне на този тип графи са т. нар. атрибутно-стойностни матрици (AVM). Те представляват вложени една в друга матрици от двойки атрибут-стойност. Всяка матрица съответства на един възел в атрибутните структури с функционални атрибути. Матрицата е маркирана с типа на възела и съдържа атрибутите, които излизат от него. Стойностите на атрибутите са (1) конкретни стойности на тип, ако възелът, до който водят, е листо в графа; или (2) матрица с двойки атрибут-стойност за възела, който е стойност на атрибутите. Тук е даден пример на такава матрица, съответстваща на горния граф, с уговорката, че матрицата започва с обекта, който има етикет *type4*:

$$\begin{array}{c}
 \begin{array}{c}
 \text{type4} \\
 \text{FEAT-A} \\
 \text{FEAT-B} \\
 \text{FEAT-C}
 \end{array}
 \begin{array}{c}
 [1] \\
 [3] \\
 [2]
 \end{array}
 \begin{array}{c}
 \begin{array}{c}
 \text{type1} \\
 \text{FEAT-A} \\
 \text{FEAT-B}
 \end{array} \\
 \begin{array}{c}
 \text{type5} \\
 \text{FEAT-A} \\
 \text{FEAT-B} \\
 \text{FEAT-C}
 \end{array}
 \end{array}
 \begin{array}{c}
 \begin{array}{c}
 [1] \\
 [2]
 \end{array} \\
 \begin{array}{c}
 [1] \\
 [3] \\
 [2]
 \end{array}
 \end{array}
 \end{array}$$

Графът се представя като пътища в самия граф, водещи до подграфи. Индексите [1], [2] и [3] осигуряват представянето на цикли в графа и/или структурно съвмествяване. Следователно може да се направи извод, че двата пътя **FEAT-A** и **FEAT-C:FEAT-A** водят до един и същи възел, т.е. съвмествят една и съща стойност.

2.2

Граматични принципи

Тук представяме най-общо идеята за граматични принципи в ОФГ и тяхната интерпретация. Граматичните принципи са логически твърдения (формули, описания), които се оценяват като верни или неверни за лингвистичните обекти. Граматичните принципи имат вида на импликации:

$$A \rightarrow B$$

където А и В са формули, които се състоят от елементарни описания (формули) и логически връзки като отрицание, конюнкция, дизюнкция, импликация. Елементарните описания определят типовете на лингвистичните обекти, последователността от атрибути, дефинирани за дадени обекти, и типовете на стойностите на атрибутите. Граматичните принципи се интерпретират по следния начин: всеки обект, който удовлетворява формулата А, трябва също така да удовлетворява и формулата В. Тук няма да навлизаме в повече детайли. Реални принципи от ОФГ ще бъдат представени в следващите раздели. Където е нужно, ще представяме формулите като атрибутно-стойностни матрици, които съдържат само атрибутите и стойностите, необходими за съответните граматични принципи. Следва един пример за импликация:

$$\left[\begin{array}{c} A \end{array} \right] \rightarrow \left[\begin{array}{c} B \end{array} \right]$$

Тук А и В са списъци от подходящи атрибути и стойности и/или типове. При представянето на конкретните принципи ще дадем конкретни стойности за тях.

2.3

Опорната фразова граматика – йерархия и принципи

След като въведохме структурата на йерархията на типовете и граматичните принципи, в този раздел ще представим основната йерархия на типовете, използвана в ОФГ, и нейните универсални⁴ граматични принципи.

Граматиката на конкретен език (например българския) в рамките на ОФГ се състои от три основни елемента: *йерархия на типовете на езиковите обекти, лексикон и принципи*.

Йерархията на типовете (вж. по-горе) е основана на отношението *тип-под-тип* и определя *онтологията* на езиковите обекти в действителността. В тази онтология обаче се дефинират не само видовете обекти, но и техните характеристики (*атрибути*). Самите характеристики се наследяват по линията на отно-

⁴ Тук под универсални принципи се разбират тези принципи на граматиката, които се приемат за валидни за всички естествени езици и представляват основата на ОФГ като лингвистична теория.

шението *тип-подтип*, т.е. характеристиките на даден тип се наследяват от всички негови подтипове, като те могат да имат и допълнителни характеристики.

Лексиконът кодира основната информация на лексикалните елементи в езика. Тази информация определя семантичната и прагматичната натовареност на лексикалната единица, както и съчетаемостта ѝ с други единици в езиковата структура.

Принципите на граматиката функционират като допълнителни ограничения върху езиковите обекти, които са допустими (възможни) според йерархията на типовете и лексикона. Именно поради факта, че принципите **ограничават** броя и типа на възможните конструкции, а не ги предписват, ОФГ се разглежда като *граматика на ограниченията*. Тъй като общият вид на ограниченията позволява езиковоспецифичните особености да се представят до голяма степен в лексикона, ОФГ се нарича още и лексикална граматика. Разделението между лексикон и принципи има лингвистичен характер. От гледна точка на формализацията те се представят с помощта на едни и същи формални методи.

ОФГ – йерархия на типовете

Формалният характер на йерархията на типовете беше обсъден на абстрактно ниво по-горе. Тук представяме реалните типове и атрибути в ОФГ, които ще използваме при моделирането на българските изречения.

Основният тип езикови обекти са знаковете – представени от типа *sign*. Знакът е общата съвкупност от фонологичния, синтактико-семантичния и прагматичния компонент, които изграждат думите или изказванията в естествения език. Най-общо се разглеждат два вида знакове: *лексикални* и *фразови* – представени съответно от типовете *дума* (*word*) и *фраза* (*phrase*). Макар че лексикалните и фразовите знакове се различават по своята структура, те съвместяват и много общи характеристики. По тази причина те се разглеждат като подтипове на типа *sign*. Всеки знак има поне две характеристики (представени като атрибути): фонологична (атрибутът **PHON**) и синтактико-семантична (атрибутът **SYNSEM**). Атрибутът **PHON** кодира фонологията на дадения знак, а атрибутът **SYNSEM** е сложна структура от обекти и техни характеристики, включваща синтактичните, семантичните и прагматичните особености на знака. Набора от синтактични характеристики ще обозначаваме като *категория на знака* (атрибутът **CATEGORY**), набора от семантични характеристики – като *съдържание на знака* (атрибутът **CONTENT**) и набора от прагматични характеристики – като *контекст на знака* (атрибутът **CONTEXT**). В рамките на *категорията* (атрибутът **CATEGORY**) има две важни характеристики (отново представени като атрибути): *опора* (атрибутът **HEAD**) и *валентност* (атрибутът **VAL** вм. *valency*). Атрибутът **HEAD** играе ролята на обобщена част на речта и определя основните граматични особености на знака. Тъй като тази характеристика е дефинирана

както за лексикалните знакове, така и за фразовите, то не само думите, но и фразите принадлежат на дадена обобщена част на речта. Структурата на характеристиката **HEAD** е доста по-сложна от атомарната характеристика за част на речта. Характеристиката *валентност* (атрибута **VAL**) определя реализираните в конкретното изказване зависими елементи⁵ (комплементи (атрибута **COMPS**) и подлог (атрибута **SUBJ**)) на дадения знак. Потенциалните зависими елементи се представят само в лексикона, т.е. дефинирани са само за лексикалните знакове чрез характеристиката *аргументна структура* (атрибутът **ARG-ST** вм. *argument structure*). В аргументната структура се записват всички аргументи на дадения лексикален знак, включително и подлогът за глаголите. В същото време елементите от *аргументната структура* се съпоставят с елементите на валентните списъци посредством различни принципи. Както беше споменато по-горе, има два валентни списъка. По този начин се моделират реализациите на различните валентности на една и съща лексикална опора. Следва матрица, показваща аргументната структура и валентните списъци на преходен глагол от типа на *чета*, както и връзката между тях:

<i>word</i>					
ARG-ST	<[1] <i>NPnom</i> , [2] <i>NPacc</i> >				
VAL	<table> <tr> <td>SUBJ</td><td><[1]></td></tr> <tr> <td>COMPS</td><td><[2]></td></tr> </table>	SUBJ	<[1]>	COMPS	<[2]>
SUBJ	<[1]>				
COMPS	<[2]>				

В тази структура *NPnom* и *NPacc* са маркери за по-детайлни атрибути структури, които дефинират синтактико-семантичните характеристики (стойността на атрибута **SYNSEM**) на именни елементи, първият от които е в именителен падеж (*nominative*), а вторият е във винителен падеж (*accusative*). Чрез индексите [1] и [2] се осъществява реализацията на аргументите в изречение като *момчето чете книга*. Ако някой от аргументите не се реализира, той просто не се копира на съответния валентен лист. Например, следните матрици показват структурите за изреченията *чете книга* и *чете*.

<i>word</i>					
ARG-ST	<[1] <i>NPnom</i> , [2] <i>NPacc</i> >				
VAL	<table> <tr> <td>SUBJ</td><td>< ></td></tr> <tr> <td>COMPS</td><td><[2]></td></tr> </table>	SUBJ	< >	COMPS	<[2]>
SUBJ	< >				
COMPS	<[2]>				

<i>word</i>					
ARG-ST	<[1] <i>NPnom</i> , [2] <i>NPacc</i> >				
VAL	<table> <tr> <td>SUBJ</td><td>< ></td></tr> <tr> <td>COMPS</td><td>< ></td></tr> </table>	SUBJ	< >	COMPS	< >
SUBJ	< >				
COMPS	< >				

⁵ Зависимите елементи (или още депendentите) на една опора могат да бъдат аргументи и адюнкти. Аргументи са съответно комплементите и подлогът.

В първия случай има реализиране на комплемента *книга*, който е именна фраза в позиция на винителна падежна форма. Във втория случай както подлогът, така и комплементът остават нереализирани на синтактично ниво. Различните съпоставяния между аргументната и валентната структура се определят чрез специални правила за съпоставяне, които няма да дискутираме тук.

Фразовите знакове притежават една допълнителна характеристика – *конституентна структура* (задава се като стойност на атрибута **DTRS** в м. daughters). Тя дефинира синтактичните съставки (конституенти) на фразата (ще ги наричаме още и *дъщери*). Стойността на характеристиката **DTRS** определя вида на фразата. На най-високо ниво се различават два вида фрази: *опорни фрази* (*headed-phrase*) и *неопорни фрази* (*non-headed-phrase*). Неопорните фрази представляват различните видове координация. От опорните фрази ще споменем фразите от типа: *опора-комплемент* (*head-complement*), представлящи отношението между опората (опорната дъщеря) и нейните комплемементи (комплементни дъщери); *опора-адюнкт* (*head-adjunct*), представляващи отношението между опората и един неин адюнкт (адюнктна дъщеря); *опора-подлог* (*head-subject*), представляващи отношението между опората и подлога (подложна дъщеря); *опора-запълваща фраза*⁶ (*head-filler*), представляващи отношението между опората и екстрактната, или още извлечената, от опората фраза (запълваща дъщеря).

Тук ще представим онези части от йерархията на типовете, които сме обяснили до момента. Първо йерархията на знаковете:

sign

PHON : *phonlist*

SYNSEM : *synsem*

word

ARG-ST : *list(synsem)*

phrase

DTRS : *con-struc*

Трябва да споменем, че *phonlist* е списък с фонологичните думи, така както са в словоредната реализация на съответния знак – фраза или дума. Типът *synsem* представлява синтактико-семантичните характеристики на знака. Аргументната структура се задава като списък от такива синтактико-семантични характеристики. Типът *con-struc* определя конституентната структура на фразата.

Различните видове фрази, споменати по-горе, се дефинират на базата на йерархията от подтипове на типа *con-struc* (съкратено от *constituent structure*), който представлява конституентната структура на знака. Тук представяме тази част от йерархията на типовете:

⁶ Запълваща е фразата, която се реализира нелокално (дистантно) спрямо своята опора.

con-struct

headed-phrase

HEAD-DTR : *sign*

COMP-DTRS : *list-of-phrases*

head-complement

HEAD-DTR : *word*

COMP-DTRS : *non-empty-list-of-phrases*

head-subject

HEAD-DTR : *phrase*

SUBJ-DTR : *phrase*

COMP-DTRS : *empty-list*

head-adjunct

HEAD-DTR : *phrase*

ADJUNCT-DTR : *phrase*

COMP-DTRS : *empty-list*

head-filler

HEAD-DTR : *phrase*

FILLER-DTR : *phrase*

COMP-DTRS : *empty-list*

non-headed-phrase

coordination-phrase

CONJ-DTRS : *set(sign)*

CONJUNCTION-DTR : *word*

Всяка фраза от тип *headed-phrase* има една опорна дъщеря (представена като стойност на атрибута **HEAD-DTR**), която е от тип *sign* и комплементни дъщери (представени като стойност на атрибута **COMP-DTRS**), които са списък от фрази. Тези два атрибута се наследяват от съответните подтипове. Например фразите от тип *head-complement* наследяват двата атрибута, но имат допълнително условие опорната дъщеря да бъде лексикална (тип *word*) и условие комплементите да са непразен списък, тоест фразите от този тип да имат поне един комплемент. Всички други подтипове на типа *headed-phrase* наследяват двата атрибута с допълнителни ограничения върху стойностите им: атрибутът **HEAD-DTR** има стойност, която е фраза, и атрибутът **COMP-DTRS** има за стойност празен списък, което е необходимо, тъй като тези фрази нямат комплементи. Но атрибутът е необходим, за да се удовлетвори Валентният принцип (вж. по-нататък). Всеки от тези подтипове има деклариран по един допълнителен атрибут за съответната неопорна дъщеря. Те са ограничени да имат стойности, които са фрази. Фразите без опора (*non-headed-phrase*) имат само един подтип – *coordination-phrase*. Той съответства на координационните фрази и има два атрибута: **CONJ-DTRS** със стойност множество от знакове (*set(sign)*) и **CONJUNCTION-DTR** със стойност лексикален знак (*word*).

Друга важна част от йерархията на типовете е подйерархията, която започва с типа *head*. Тази част от йерархията отговаря за дефинирането на граматичните характеристики на знака, които се разглеждат като обобщена част на речта. Основна роля играе и атрибутът **MOD** (съкратено от *modifier*, тоест модификатор), който изразява отношението между адюнкта и неговата опора. Действието е описано при граматичните принципи по-нататък. Следва съответната част от йерархията на типовете:

```

head
  substantive (subst)
    PRD : boolean
    MOD : mod-synsem
  noun
    CASE : case
  verb
    VFORM : vform
    AUX : boolean
    INV : boolean
  adj
  adv
  prep
    PFORM : pform

```

Тук представяме само основните части на речта. Те определят основните видове фрази като именна фраза (*noun phrase – NP*), глаголна фраза (*verb phrase – VP*), фраза на прилагателното (*adjectival phrase – AP*), фраза на наречието (*adverbial phrase – AdvP*) и предложна фраза (*prepositional phrase – PP*). Съответните типове наследяват двата атрибута на типа *subst*: **PRD** (от *predicative*, предикативен), който определя дали фразата е използвана предикативно, или не, и **MOD**, който определя селективния потенциал на адюнктите.

Другите части на ОФГ йерархията от типове не играят такава важна роля за принципите, които представяме тук, и затова ще ги пропуснем. Повече информация може да се намери в (Pollard and Sag 1994).

ОФГ – основни граматични принципи

В този раздел представяме главните граматични принципи на ОФГ. Те представляват основата на всяка ОФГ граматика за даден естествен език.

Принцип на предаване на опорните характеристики
(*Head Feature Principle (HFP)*):

В една опорна фраза (от тип headed-phrase) стойността на HEAD характеристиката на фразата съпада със стойността на HEAD характеристиката на опорната ѝ дъщеря.

Най-общо този принцип изразява лингвистичната интуиция, че в една фраза с опора (опора-комплемент, опора-адюнкт, опора-подлог и т.н.) опорната дъщеря определя обобщената част на речта за цялата фраза. Представяме го и като правило:

$$\left[\begin{array}{c} \text{headed-phrase} \end{array} \right] \rightarrow \left[\begin{array}{cc} \text{SYNSEM} | \dots | \text{HEAD} & [1] \\ \text{DTRS} | \text{HEAD-DTR} | \text{SYNSEM} | \dots | \text{HEAD} & [1] \end{array} \right]$$

Както се вижда, в описанията даваме само най-необходимата информация. Последователностите от атрибути като **SYNSEM** | ... | **HEAD** дефинират път в съответните граfi. Следователно принципът ни казва, че ако даден обект (възел в граф) е от тип *headed-phrase*, то двата пътя имат една и съща стойност (или, с други думи, стигат до един и същ възел).

Валентен принцип

(*Valence Principle (VALP)*):

В една опорна фраза стойността на всеки валентен атрибут VF (според йерархията по-горе SUBJ и COMP) съвпада със стойността на съответния валентен атрибут VF на опорната ѝ дъщеря, като се изключат synsem обектите, съответстващи на стойностите на атрибутите SYNSEM на реализираните зависими дъщери.

Валентният принцип управлява действителното реализиране на аргументите на лексикалната опора. Първо, всички комплементи (ако има такива) се реализират. Това следва от Валентния принцип и йерархията на типовете, която казва, че във фраза от типа *опора-комплемент* опорната дъщеря е лексикална. След това се реализира подлогът (ако има такъв). Следва правилото за реализация на комплементите. Те се избират наведнъж от опората. За да бъдат подходящи за изискванията на опората, техните синтактико-семантични характеристики трябва да съвпадат с изброените характеристики на комплементния списък. В правилото това съвпадение схематично е означено чрез индекса [1], макар че по принцип трябва да изразим равенство на списъци⁷. Тъй като комплементите се реализират едновременно, след тяхната реализация към цялата фраза се предава празен списък. Следователно опората не изисква повече комплементи. Аналогично е правилото за реализация на подлога.

$$\left[\begin{array}{c} \text{head-complement} \end{array} \right] \rightarrow \left[\begin{array}{cc} \text{SYNSEM} | \dots | \text{VAL} | \text{COMPS} & <> \\ \text{DTRS} | \text{COMP-DTRS} | \dots | \text{SYNSEM} & [1] \\ \text{DTRS} | \text{HEAD-DTR} | \dots | \text{VAL} | \text{COMPS} & [1] \end{array} \right]$$

⁷ Това опростяване не ощетява по никакъв начин адекватността на обяснението за работата на принципа.

Реализирането на адюнктите се управлява от следния принцип:

Принцип на адюнкта

(*Head-Adjunct Principle (AdjunctP)*):

Във всяка фраза от типа опора-адюнкт стойността на атрибута MOD на адюнктната дъщеря съпада със стойността на атрибута SYNSEM на опорната дъщеря.

Този принцип показва, че адюнктната дъщеря избира опората си чрез своя **MOD** атрибут. Атрибутът ‘вижда’ синтактико-семантичните характеристики на опората. Като правило принципът изглежда по следния начин:

$$\left[\begin{array}{c} \text{head-adjunct} \end{array} \right] \rightarrow \left[\begin{array}{cc} \text{DTRS|HEAD-DTR|SYNSEM} & [1] \\ \text{DTRS|ADJUNCT-DTR|SYNSEM|...|MOD} & [1] \end{array} \right]$$

Равенството на атрибута **MOD** на адюнкта и на атрибута **SYNSEM** на опората означава, че адюнктът може да наложи изисквания към синтактико-семантичните характеристики на опората.

Принцип на семантиката

(*Semantic Principle (SemP)*):

В една опорна фраза стойността на атрибута CONTENT на фразата съпада със стойността на атрибута CONTENT на семантичната ѝ опора.

Този принцип управлява семантичната интерпретация на една фраза на базата на семантиката на дъщерите. Семантичната опора на дадена опорна фраза е адюнктът, ако фразата е от тип опора-адюнкт, а в другите случаи е опората.

Както ще видим в раздела, посветен на прагматичните фрази, по-нататък в книгата ние променяме тази дефиниция за семантична опора, като разделяме фразите опора-адюнкт на два подтипа според вида на адюнкта – прагматичен или семантичен. При прагматичните адюнкти семантичната опора съпада със синтактичната опора. Тук няма да даваме правилото.

Последният принцип, който ще представим, се грижи за реализирането на запълващата фраза при фрази от тип *head-filler*:

Принцип на запълващата фраза

(*Head-Filler Principle (FillerP)*):

За всяка фраза с атрибут DTRS, който има стойност от тип head-filler, стойността на атрибута SLASH на опорната дъщеря съдържа стойността на атрибута LOCAL на запълващата дъщеря. Стойността на атрибута SLASH за цялата

фраза съпада със стойността на атрибута SLASH на опорната дъщеря минус стойността на атрибута LOCAL на запълващата дъщеря.

Тук използваме два атрибута, които не са въведени по-горе. Те се дефинират в рамките на синтактико-семантичните характеристики на даден знак. Атрибутът **LOCAL** свързва знака с локално дефинираната информация като категория, семантика и т.н. Атрибутът **SLASH** дефинира списък, в който се натрупва информация за нереализираните локално дъщери. Атрибутът **SLASH** на дадена фраза се предава на следващата фраза в йерархията. Когато се появи някоя от нереализираните локално дъщери, информацията за нея се изтрива от списъка на атрибута **SLASH**, преди да бъде предаден към следващата фраза.

Така представените принципи са само част от реалните принципи в една ОФГ граматика. Към тях трябва да се добавят принципи за словореда, за съгласуването, за интерпретация на кванторите в изречението и много други. Една съществена част от ОФГ граматиката е лексиконът, състоящ се от така наречените лексикални принципи, които задават различни характеристики на думите. Формално те също се изразяват като импликации.

Представената архитектура на една ОФГ граматика има за цел да демонстрира механизмите на описание на формалната граматика на даден естествен език. Първо йерархията на типовете определя основните типове лингвистични обекти и техните характеристики (атрибути). На базата на йерархията на типовете могат да се построят и потенциалните лингвистични обекти. Действителните лингвистични обекти се определят с помощта на граматичните принципи, които задават различни отношения между атрибутите и техните стойности на лингвистичните обекти.

В тази книга се ръководим от горните принципи при представянето на различните типове фрази и явления в българския език. Където е нужно, въвеждаме нови принципи, за да представим адекватно езиковия материал. Тъй като представянето на езиковите обекти под формата на атрибутни графи е прекалено сложно, ние се абстрахираме от множество детайли, които биха играли важна роля при имплементацията на една такава граматика, но тук могат да бъдат заместени с по-обща категории. Затова графичното представяне на езиковия материал ще отразява само онези свойства на езиковите елементи, които играят основна роля при съответните анализи. Това опростяване на формалния апарат е представено в следващата глава.

Представяне на формалния модел на ОФГ в графичен вид

Формалният апарат на ОФГ е много мощно моделиращо средство, но в същото време и много тежко за възприемане и описание. За да можем да представим анализите си в по-разбираем и четивен вид, ние проектираме сложната структура на езиковите елементи от типа *sign* в категории, които обобщават всички характеристики, дефинирани по-горе.

Характеристиките, които определят структурата и поведението на различните фрази, са следните:

- **Конституентност.** Отношението фраза-конституенти, представено чрез атрибута **DTRS**, е основно за структурата на всяка фраза. Това отношение се представя чрез специален вид насочени графи, наречени дървета.
- **Категория.** Всеки знак се представя като възел на даденото дърво. Категорията на знака се представя като етикет на дадения възел. Както ще видим по-нататък, този етикет може да бъде разширяван с допълнителна информация при необходимост. Основните категории са *VP*, *NP*, *AP*, *PP*, *AdvP*.
- **Депendentност.** Етикетите на възлите в дървото отразяват и отношението между опората и зависимите части. Например *NPA*, където *A* обозначава фраза от тип опора-адюнкт, или *VPC*, където *C* обозначава фраза от тип опора-комплемент.
- **Словоред.** ОФГ отделя конституентната структура на знака от словоредата на конституентите в тази структура. При представянето на анализите ще се придържаме към оригиналния словоред на дадения пример.
- **Кореферентни отношения**⁸. Няма да можем да отделим място на семантиката, която играе важна роля за определяне на вярната структура на даден езиков обект. Но едно явление, което не може да бъде оставено непредставено, е кореферентията между отделни конституенти в рамките на едно или няколко изречения. Тези отношения в графичен вид се представят чрез дъги, които свързват съответните възли на дървото.
- **Неизразени конституенти.** Тук се имат предвид неизразеният подлог и елипсите. Неизразеният подлог много често участва в кореферентни отношения с конституенти на изречението. Елипсите са представени заедно с връзките им към съответните възстановяващи части на изречението.

⁸ За кореферентно отношение смятаме връзката между поне два елемента, които сочат към един и същ обект от действителността.

- **Морфологично-синтактични характеристики.** Тези характеристики са важни за правилното моделиране на много явления. Те няма да присъстват в графичното представяне на анализите, но ще бъдат коментирани на съответните места в текста. Тези характеристики се предават нагоре по дървото вследствие на действието на Принципа на опората.

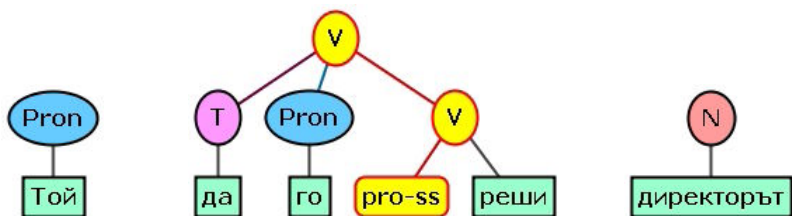
На базата на тези характеристики възлите в дървото се разделят на следните групи:

- **Лексикални възли.** Тези възли съответстват на знакове от тип *word* (дума). Те включват не само лексикалните елементи, представени в лексикона, но и аналитичните форми на думите, образувани чрез прилагане на определени лексикални правила. Някои от категориите на такива възли са: *N*, *V*, *Prep* съответно за съществително, глагол, предлог. Самите думи ще бъдат представени като отделни възли под съответните лексикални възли.
- **Фразови възли.** Тези възли отразяват както йерархията на фразите (като опора-комплемент (*head-complement*), опора-адюнкт (*head-adjunct*)), така и съответните синтактични области (като глаголни фрази, именни фрази и т.н.). Категориите за тези възли са комбинация от двата типа информация. Например, *VPA* означава глаголна фраза от тип опора-адюнкт (идва от *VPA(djunct)*) или *NPC* означава именна фраза от тип опора-комплемент (идва от *NPC(omplement)*).
- **Функционални възли.** Това са спомагателни възли, които дават възможност да се представят важни характеристики на конституентите, над които спомагателните възли доминират. Такива възли имат за етикети категории, които представят тези характеристики като различни типове подчинени изречения (клаузи), например *CLDA* за подчинени изречения с да-конструкция, *Pragmatic* за прагматични елементи в рамките на даден конституент и т.н.
- **Текстови възли.** Текстовите възли са листата на всяко дърво и представляват последователностите, които съответстват на думите и пунктуацията в оригиналното изречение.

Всеки един от тези видове възли е описан детайлно в следващите раздели на книгата. Тук само показваме как те се изобразяват графично.

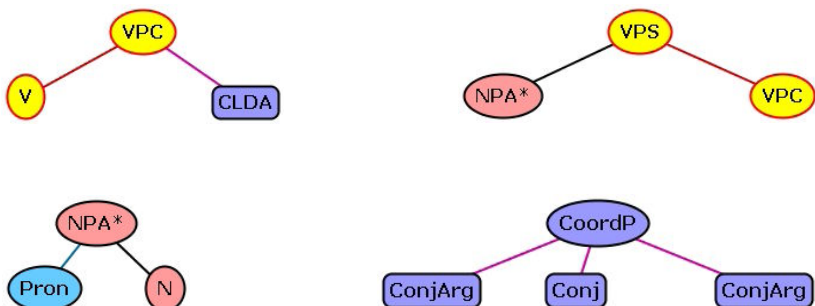
Лексикалните и фразовите възли се представят графично като овали, а техните етикети се разполагат вътре в овала. Лексикалните възли се различават от фразовите възли по техните етикети. Допълнително овалите са оцветени в различни цветове в зависимост от категориите си. По този начин читателят може по-лесно да проследява различните пътища в дърветата. Опората на всяка фраза се извежда на базата на типа фраза. Типът на зависимите части за всеки вид фраза ще бъде представен при описанието ѝ. За повече подробности вж. следващите раздели. Следните схеми показват няколко лексикални и фразови възела:

Лексикални възли:



Тук са изобразени няколко лексикални възела: два *Pron* възела, един *T* възел, два *V* възела и един *N* възел. Както може да се забележи, освен единични, лексикалните елементи могат да бъдат комплексни, вложени структури. В този пример единият от глаголните елементи представлява аналитична глаголна форма с клитики, която се състои от спомагателната частица *да* (етикет *T*), прономинална клитика (етикет *Pron*) и глагол (етикет *V*). Вложеният възел има като дъщеря също и маркера за неизразен подлог *pro-ss*. Лексикалните възли имат като наследници своите словоформи.

Фразови възли:



Наследниците на даден фразов възел съответстват на структурата му. Наследниците могат да бъдат други фразови възли, лексикални възли или функционални елементи. На горната фигура има следните елементи: *VPC* възел, който представлява глаголна фраза от тип *опора-комплемент*; лексикален възел *V* с комплемент *CLDA*; *VPS* възел, който представлява глаголна фраза от тип *опора-подлог* с опора *VPC* и подлог *NPA*; *NPA* възел за именна фраза от тип *опора-адюнкт* с опора лексикален елемент *N* и адюнкт – друг лексикален елемент *Pron*. Звездичката в етикета обозначава, че този елемент е аргумент на корелативна или друга релация към друг елемент в цялостната структура, която не е показана тук; *CoordP* възел, който представлява координационна фраза, няма опора и има като деца функционални възли за координационните аргументи и съюзите.

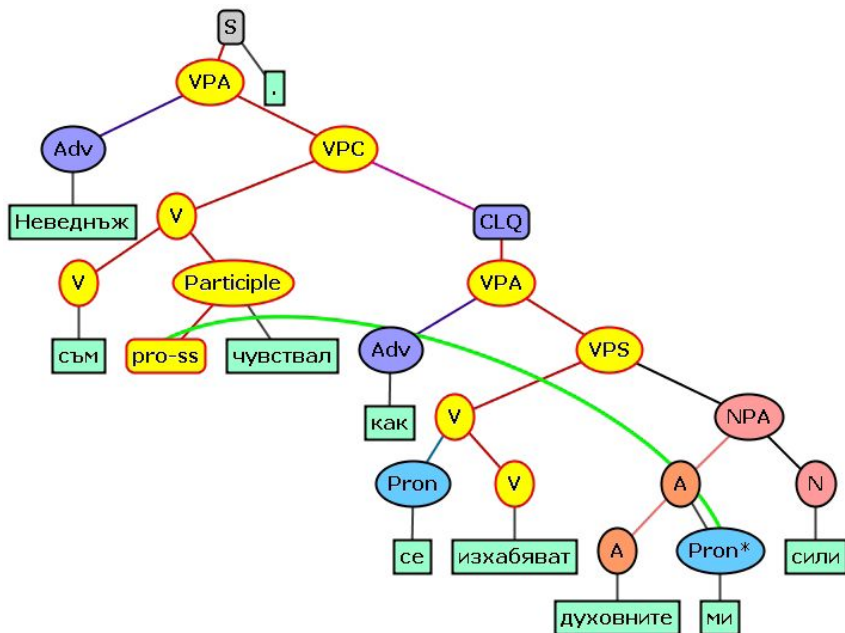
Функционалните възли се представят графично като правоъгълници със заоблени върхове, а техните етикети се разполагат вътре в правоъгълниците. Функционалните възли имат само по един наследник – възел, на който те приписват дадена характеристика. Понякога препинателните знаци също са наследници на функционален възел. Това най-често е вярно за корена на дървото.

Функционални възли:



В тези примери има: възел *S* за корен на изречението с наследници - възел *CoordP*, обозначаващ координационна фраза, и точката в края на изречението; възел *CLDA*, който представя подчинено изречение с да-конструкция. В този случай изречението е самата да-конструкция; възел *ConjArg*, който представлява аргументът на една координационна фраза; и възел *Conj*, който представлява съюзът на една координационна фраза заедно със задължителната запетайка пред него.

Текстовите възли се представят като правоъгълници със съответния текст в тях. Както беше споменато по-горе, те представляват реалните думи в изречението, което се анализира, и включват както словоформите, така и пунктуацията. Те винаги са листата на дървото. Примери на текстови възли са дадени в следното дърво за изречението *‘Неведнъж съм чувствал как се изхабяват духовните ми сили’*:



Текстовите възли в дървото (т.е. листата му без елемента *pro-ss*) представят деветте думи в изречението и крайната точка (.). Над всяка дума има лексикален възел с подходящ етикет, който зависи от морфологично-синтактичните характеристики на думата. Има три сложни лексикални възела – два глагола и едно прилагателно – *съм чувствал*, *се изхавяват* и *духовните ми* (те са обяснени детайлно в съответния раздел по-нататък). Дървото съдържа пет фразови възела: *NPA*, *VPS*, *VPA*, *VPC* и *VPA*. Именната фраза (*NPA*) се състои от съществителното *сили* и сложната прилагателна форма *духовните ми*. Съществителното е опора, а прилагателното е адюнкт. Това е отразено в етикета на възела. Следващата фраза е глаголна фраза от тип опора-подлог – *VPS*, където глаголът *се изхавяват* е опората, а *NPA* е подлог. Въпросителното местоимение *как* е маркирано като наречие (*Adv*), тъй като играе тази роля в изречението⁹. То се явява адюнкт на опората *VPS* и двете заедно образуват глаголна фраза *VPA*. Фразата *VPA* е подчинено въпросително изречение. Този факт се маркира с функционалния възел *CLQ*. Това подчинено изречение е комплемент на глагола *съм чувствал* и двете образуват глаголна фраза от тип *опора-комплемент* (*VPC*). Наречието *неведнъж* е адюнкт към *VPC* и образуват глаголна фраза от тип *опора-адюнкт* *VPA*. Функционалният възел *S* маркира корена на дървото. Към него е прикрепена и крайната точка на изречението. В дървото е отразено и кореферентното отношение между подлога на основното изречение и притежателната клитика *ми* в *духовните ми*. В този случай неизразеният подлог на главното изречение е маркиран с функционалния елемент *pro-ss*. Лексикалният елемент над клитиката е маркиран допълнително със звездичка за по-голяма прегледност.

Следващият важен въпрос с оглед на графичното представяне на синтактичните анализи е отношението между линейната наредба на възлите на дървото и конституентната структура на фразите, които се обозначават чрез тези възли. В ОФГ лингвистичното знание за конституентната структура на фразата е отделено от лингвистичното знание за наредбата на конституентите в изречението. Реализацията на различните зависими части в конституентната структура на изречението в ОФГ се управлява чрез множество от схеми за непосредствена доминация. Те специфицират за даден вид фраза какъв е типът на опората и на зависимата част. Според стандартните схеми за непосредствена доминация в ОФГ всяка фраза от тип *опора-комплемент* трябва да има лексикална опора и съответните комплементи да са фрази. Следствието от тази схема е, че комплементите се реализират първи, преди всички други зависими части. Приемаме, че след комплементите (ако ги има) се реализира подлогът (ако го има явно представен) и след него се реализират адюнктите. Следователно, реализацията на зависимите части следва йерархията: *комплемент* → *подлог* → *адюнкт*. Тази йерархия определя някои от решенията, избрани от нас при анализа на определени видове фрази и явления. Ако една фраза е от вида *VPA*, това означава, че или комплементите и подлогът са вече реализирани, или те

⁹ Морфологично-синтактичните му характеристики посочват местоимения му произход.

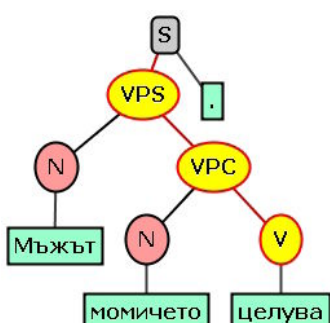
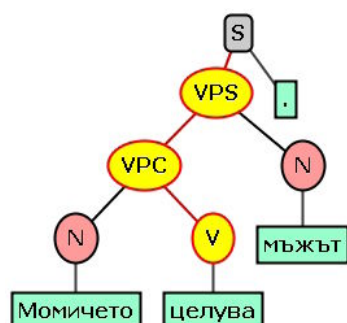
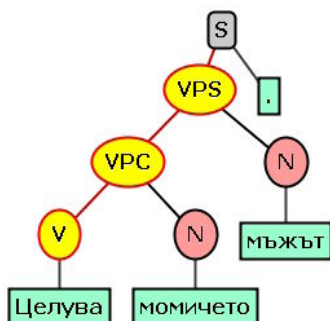
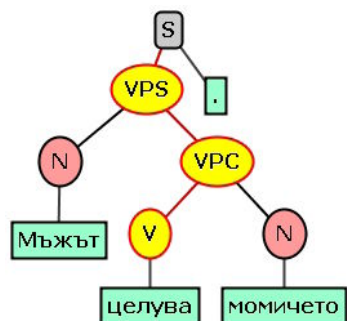


не са явно изразени, или лексикалната опора няма някои от тях. Възможна е реализацията и на други адюнкти. Ако фразата е от вида *VPS*, това означава, че фразата няма адюнкти, а комплементите ѝ са вече реализирани, не са изразени явно или липсват в аргументната структура на опората. Аналогично, *VPC* на най-високо ниво означава, че подлогът не е изразен и няма адюнкти. Забележете, че всяка една от трите типа фрази (*VPC*, *VPS* и *VPA*) може да липсва в анализа на дадено изречение. Следвайки горната реализация на зависимите части, ние казваме, че подлогът е по-висока зависима част от комплементите и адюнктът е по-висока зависима част от подлога и комплементите. В обратна посока ще казваме, че комплементът е по-ниска зависима част от подлога и адюнкта и че подлогът е по-ниска зависима част от адюнкта. Тези дефиниции са в сила само когато комплементите, подлогът и адюнктите имат обща лексикална опора.

На тази база нашите допускания са:

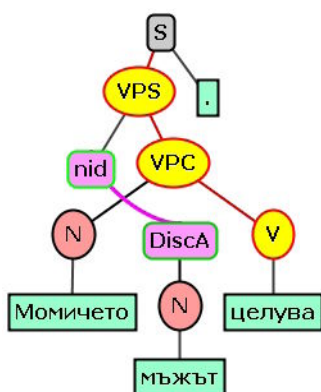
- Конституентната структура е отделена от линейната подредба на конституентите.
- Елементите на конституентната структура нямат предварителна подредба. Тяхната линейна подредба се задава с допълнителни правила на граматиката или в лексикона, когато линейната подредба е специфична за даден клас от лексикални елементи.
- Елементите на конституентната структура се реализират обикновено непосредствено един до друг. Различната им подредба не се отбелязва по определен начин. Тоест всички допустими подредби се смятат за еднакви конституентни структури. Например следните подредби се приемат за една и съща конституентна структура: *‘Мъжът целува момичето’*, *‘Целува момичето мъжът’*, *‘Момичето целува мъжът’*, *‘Мъжът момичето целува’*. Различната линейна подредба има, разбира се, своята роля при интерпретацията на даденото изречение, но тя не се изразява чрез конституентната структура.
- Всеки конституент изразява граматична функция.
- Категорията на опората на фразата определя областта на действие на фразата.

Следната диаграма показва графичното представяне на четирите изречения:



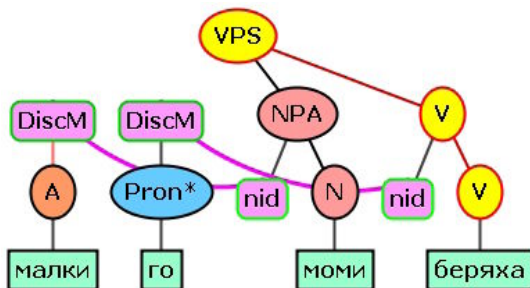
Ако елементите на една фраза не са разположени непосредствено един до друг, се появяват така наречените *дистантно разположени конституенти*. Разграничаваме три вида дистантно разположени конституенти в зависимост от свойствата на опората.

Смесване на зависимите части на една и съща опора (обозначава се с функционалния елемент *DiscA*). Това е случаят, когато една по-висока зависима част е реализирана между опората и по-ниска зависима част или между две по-ниски зависими части. Тоест имаме опора и поне две нейни зависими части от различен вид, които са разположени заедно, но в разменени позиции. Ето словореден вариант на горното изречение, при което подлогът е разположен между комплемента и опората:



За да представим този случай графично, ние маркираме подлога с функционалния елемент *DiscA* и го свързваме с друг функционален елемент *nid* (от *non immediate dominance*), който показва принадлежността на съответния елемент към съответната структура. Елементът *nid* показва само принадлежността към конституентната структура на съответната фраза, но не показва словоредната позиция.

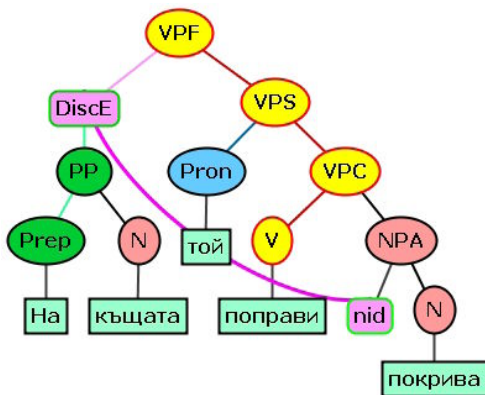
Смесване на зависимите части на две опори (обозначава се с функционалния елемент *DiscM*). Това е случаят, когато две независими една от друга опори (тоест никоя от тях не е в рамките на зависима част на другата опора) имат поне по една зависима част. Освен това опорите и зависимите им части са подредени така, че никоя от зависимите части не е непосредствено до своята опора. В този случай зависимите части са маркирани чрез функционалните елементи *DiscM*, които са свързани със съответните *nid* елементи при съответните опори. Такива примери са редки. Затова примерът е от народното творчество:



Външно позициониране на вътрешен конституент (обозначава се с функционалния елемент *DiscE*). Това е случаят, когато зависима част от една фраза е реализирана в позиция, която е в рамките на друга по-висока фраза. В следния пример *PP* адюнктът на опората *покрива* е разположен дистантно от управляващата го именна фраза. Дистантно разположеният възел е маркиран с елемента *DiscE*, който винаги е наследник на възел от тип *VPF* (глаголна фраза със запълващ елемент).

Тук още по-ясно личи, че възелът *nid* показва принадлежност към конституентната структура на изречението, но не и словоредни зависимости.

В следващите раздели ще представим различните типове фрази.



В ОФГ лексикалните елементи отговарят на лексикалните знакове¹⁰ (т.е. тип *дума* (*word*)). Невинаги отношението лексикален елемент : ортографична дума е 1:1. Също така, лексикалните елементи се различават от лексемите в смисъла на инвариант на словоформите на дадена дума. Те всъщност представляват словоформите на лексемата, но включвайки два основни компонента: наследената информация от лексемата и добавените техни специфични характеристики. Например лексикалният елемент *столове* наследява информацията за съществително, мъжки род от *стол*, като добавя информацията за множествено число. За определянето на даден езиков знак/знакове като лексикален роля играят различни фактори, например единното морфологично значение, типът на компонентите, непротиворечивостта с анализите на другите езикови нива. Лексикалните елементи могат да включват: самостоятелни словоформи, последователност от няколко ортографични думи (т. нар. *multiwords*), както и аналитични форми на думите. Напр. наречието за логическо уточняване *може би* се състои от две ортографични думи – *може* и *би*. Но само заедно те образуват значението на определен тип наречие. Понякога лексикалните елементи съдържат компоненти, които са по-скоро прозодична, отколкото морфологична или смислова цялост (в случаите с клитиките).

Характерна черта на нашия анализ е фактът, че приемаме съществуването на стратифицирани лексикални елементи, т.е. те могат да бъдат на няколко нива. По-точно казано, комбинацията от лексикални елементи може да образува пак лексикален елемент.

Нека разгледаме типовете лексикални елементи. Те се съотнасят с частите на речта¹¹. Това са: глагол (*V*), причастие (*Participle*), деепричастие (*Gerund*), съществително име (*N*), местоимение (*Pron*), прилагателно (*A*), числително (*M*), фамилно име или прилагателно от собствено име (*H*), наречие (*Adv*), съюз (*C*), предлог (*Prep*), частица (*T*) и междуметие (*I*).

Трябва да отбележим, че лексикалното ниво е мястото, където си взаимодействат морфологичното и синтактичното ниво. Тук става преход от морфологич-

¹⁰ В по-новите разработки по ОФГ има и други схващания за йерархията на езиковите знакове. Напр. в (Sag et. al 2003, p. 229) *знакът* (*sign*) се заменя с *израз* (*expression*). Йерархията се реструктурира така, че на най-високото ниво е *синтактико-семантичният обект* (*synsem*), под него има две дъщери: *израз* (*expression*) и *лексема* (*lexeme*). От своя страна *изразът* (*expression*) се подразделя на *дума* (*word*) и *фраза* (*phrase*). Но в тази книга ние тенденциозно се придържаме към традиционните за теорията постулати, базирани на (Pollard and Sag 1994).

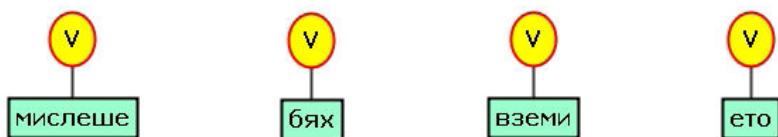
¹¹ Трябва да отбележим, че нашият списък с части на речта е по-разширен в сравнение с традиционното деление на 10 части на речта.

но-синтактичните характеристики на думите към синтактичните категории. В традицията на компютърната лингвистика представяме набора от морфологично-синтактични характеристики като множество от морфологични етикети (анотационни маркери, англ. tags). Списъкът с етикетите, които сме използвали в рамките на проекта BulTreeBank¹², е представен като приложение в края на книгата. А тук, в текста, за всяка категория лексикални знакове даваме кратки въведения за връзката им с морфологичните етикети. Описанията се състоят от букви, които са съкращения на дадено значение, и специалните символи процент (%) и диез (#). Всяка буква се съпоставя на себе си в морфологичния етикет, като обозначава съкратено някаква езикова характеристика. Символът % се съпоставя на точно една буква (която и да е), а символът # се съпоставя на произволна (възможно празна) последователност от букви. По този начин описанието «V%%f#» например, обобщава всеки морфологичен етикет за лична глаголна форма. Част от информацията, маркирана чрез процент и диез, остава скрита. По-подробно за принципите и целите на създаване на множества от морфологични етикети вж. (Паскалева 2007).

4.1

Лексикален елемент глагол

Лексикалният елемент глагол (V) се отнася до самостоятелна глаголна форма, глаголна частица, аналитична глаголна форма (залогови форми, условно наклонение, аналитичните времена и др.) или комбинация на глагол и клитика(и). На най-ниско ниво лексикалният елемент глагол (V) отразява следните морфологично-синтактични характеристики: “V%%f#”, “V%%z#”, или “Tv”, където V означава глагол, T означава частица, f означава лична форма в извършително наклонение, v означава ‘свързан с глагол’, z означава повелително наклонение. В графичен вид тези елементи са представени, както следва:



Освен чрез лични или нелични (причастни и деепричастни) форми, глаголната опора може да бъде изразена и чрез координационна структура (да *пея и пия до насита*), вербализирана форма (*И той цоп! в локвата*) или елипса на глагол (*Вол се връзва за рогата, а човек – за езика.*). Когато лексикален глаголен елемент е модифициран от фокусиращи думи като *само*, *даже* и др., проекцията¹³ пак е лексикален елемент глагол (V). За лексикални елементи смятаме още перифрастични

¹² www.bultreebank.org

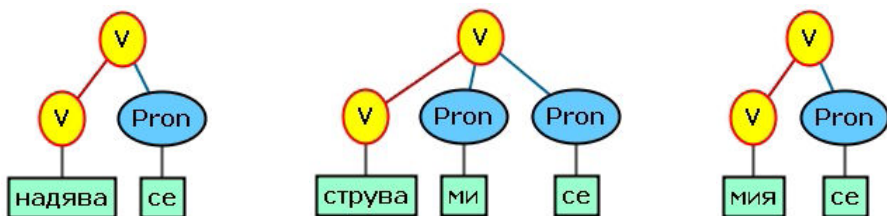
¹³ Под понятието проекция разбираме копирането на характеристики от съставките на една фраза към самата фраза. Разбира се, в случая копирането става първоначално на максималното лексикално ниво.

конструкции от типа на *имам направени*, които са смислово идентични на формата за минало неопределено време на глагола *направля*, а именно - *направил съм*.

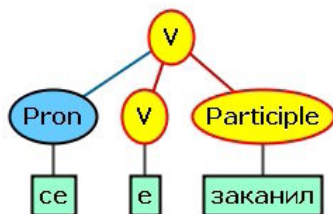
Интересно е да се проследят типовете лексикални глаголни елементи на следващото ниво на сложност, т.е. съставеното от поне два или повече елемента. Такива са: глаголни комплекси¹⁴, в които опората е финитен глагол с прилежащи към него клитики; спомагателни частици; емфатични наречия и причастия. Следват описанията на отделните типове.

Глаголен комплекс с възвратни винителни и дателни клитики

Кратките винителни (морфологичен етикет - *Ppxla*) и дателни клитики (морфологичен етикет - *Ppxld*) отразяват както истински рефлексивните, така и нерефлексивните значения на глаголите, при които възвратната по форма клитика е част от лексикалното значение на глагола (напр. глаголт *усмихвам се* не може да се разгърне като **усмихвам себе си*). Тъй като тези разлики в поведението на глаголните комплекси са кодирани в лексикона чрез лексикални спецификации и правила, ние не ги маркираме експлицитно в лексикалната структура. Ето примери:



Първият и вторият пример отразяват нерефлексивни употреби на клитиката, докато последният пример отразява (поне в едно от тълкуванията си) рефлексивна употреба в смисъла на *мия себе си, а не друг*. Разнообразието от словоредни възможности се реализира без проблем в рамките на посочения анализ.

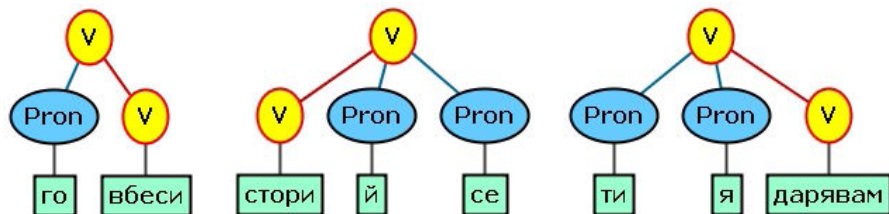


Причината да допускаме морфологична композиционалност и при нерефлексивните комплекси е възможността клитиката да не се реализира непосредствено преди или след глагола при определени условия. Напр. в горния пример двете части на глаголния комплекс *заканя се* могат да бъдат разделени от спомагателния глагол *съм* в 3 л., ед. ч., минало неопределено време.

¹⁴ Глаголният комплекс може да се отнася за различни глаголни структури. Той може да бъде напр. аналитична форма на глагола (*бях ходил, бих казал*), форма на глагола с клитики (*ядеше му се*) или да-конструкция със или без клитики – *да дойда, да му го кажа*.

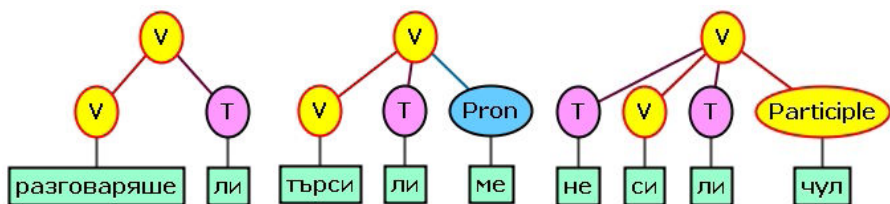
Глаголен комплекс с невъзвратни винителни и дателни клитики

Тези клитики са прякообектните (винителни) и косвенообектните (дателни) аргументи на глагола. В нашия анализ те се реализират на лексикално ниво, което означава, че задават параметрите за комплемента, но не се проектират като такива на фразово ниво. По този начин се решава проблемът с удвоените допълнения в българския език. Тъй като в синтаксиса не може да има два езикови елемента с една и съща синтактична функция, ние разпределяме приноса на клитиките и другите елементи на две равнища – лексикално и синтактично. Клитиките задават само най-общите характеристики на комплемента: лице, число, род (в 3 л., ед. ч.), преди той да бъде специфициран по-пълно чрез именна или предложна фраза. Списъкът на клитиките е, както следва: **винителни** - *ме, те, го, я, го, ни, ви, ги* и **дателни** - *ми, ти, му, ѝ, му, ни, ви, им*. Всички клитики се реализират едновременно, като дателната клитика винаги предхожда винителната. Граматичната информация, която носят клитиките, се кодира в аргументната структура на глагола (**ARG-ST**). После се изпълнява изискването истинските аргументи и клитиките да споделят едни и същи стойности на характеристиките лице, число и род. Следват примери. В първия пример има винителна невъзвратна клитика, реализирана преди глагола. Във втория пример е представен глаголен комплекс с дателна невъзвратна и възвратна по форма се клитика, като дателната предхожда възвратната. Третият пример демонстрира подреждане на дателна и винителна невъзвратни клитики преди глагола. Словоредът не е каноничен, тъй като примерите са извадени от реални изречения, чиято структура влияе на линейната подредба.

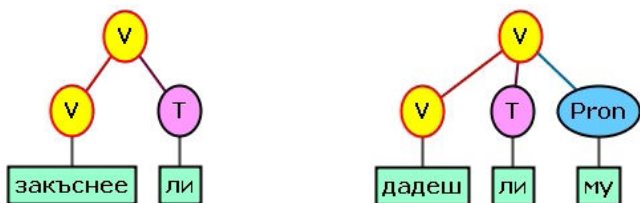


Глаголен комплекс с клитиката 'ли'

Клитиката *ли* може да изпълнява различни функции, напр. да бъде маркер за въпросителност, синоним на условния съюз *ако* и др. (вж. по-подробно в (Тишева и Джонова 2002; Тишева 2005)). Разположението на клитиката до финитна или нефинитна глаголна форма зависи от наличието на други клитики около нея. По-често се реализира в непосредствена близост до нея (пре- или постпозиция), но може да бъде и дистанцирана от друга клитика (напр. от винителната клитика в изречението: *Не си ли МЕ видял?*). Например:

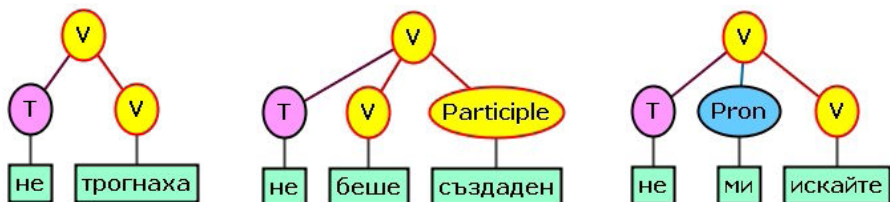


Следват примери, в които *ли* е синоним на съюза *ако*. В примерите по-долу клитиката се свързва с глаголи от свършен вид:



Глаголен комплекс с отрицателната клитика 'не'

Отрицателната частица *не* винаги е в началото на глаголния комплекс, с изключение на случаите, при които присъства и частицата *да* (вж. по-нататък). Примери:

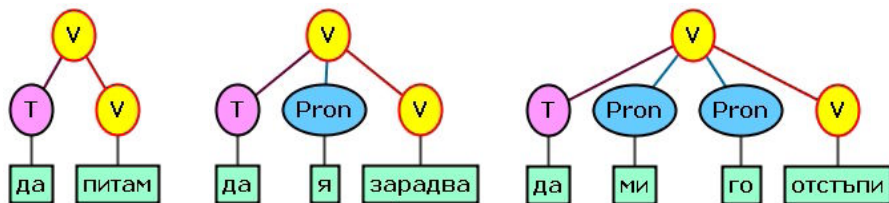


В първия случай е представен най-простият пример на комбинация от отрицателна частица и глагол. Вторият пример демонстрира по-усложнена структура на най-ниското лексикално ниво. Тук трябва да отбележим, че дърветата следват т.нар. 'плоска структура' (flat structure) на графично представяне, при която всички елементи са на едно равнище. Това е така поради възможността да се изведе морфологичната структура от нейните компоненти. В действителност обаче първо се свързват компонентите на аналитичната морфологична структура, а след това се прибавя и отрицателната частица. В горния пример морфологичната структура е причастнострадателен залог на глагола *създам* (*беше създаден*). Последният пример показва по-ясно в линеен план лявата маркираща роля на отрицателната частица в глаголния комплекс.

Глаголен комплекс със спомагателната частица 'да'

В езиковедската литература има много и различни трактовки за морфологичните етикети и функциите на формата *да*. Най-често тя се разглежда като частица и съюз (вж. за по-подробна библиография върху различните мнения и по-подробно представяне на нашите схващания в (Симов и Колковска 2004)). В съгласие с идеите на ОФГ и специално за този случай ние решихме да приемем по-скоро идеите на Конструкционната граматика (Construction grammar), отколкото на композиционалните, и да разглеждаме да-конструкцията като комплекс. Ето защо смятаме формата *да* за частица с два подтипа: утвърдителна частица (с морфологичен етикет *Ta*) и спомагателна глаголна частица (с морфологичен етикет *Tx*). Утвърдителната частица изразява съгласие. Спомагателната глаголна частица обаче има по-широки функции. От една страна, тя е част от конструкция-наследник на инфинитивната форма в старобългарския език – '*да+глагол*'. В този смисъл за нас *да* е винаги част от т. нар. 'да-конструкция'. Дори когато въвежда подчинено изречение, тя не е с роля на субординатор, а участва в 'да-конструкцията', която може да бъде аргумент или адюнкт на глагола/името в главното изречение¹⁵. От друга страна, 'да-конструкцията' има модални употреби, като напр. в *Ех, да бях млад!* В тези случаи тя въвежда условни или други модални значения и очевидно няма връзка с инфинитива. От гледна точка на морфологичните особености на глаголния комплекс спомагателната частица *да* в типичната си употреба изисква глагол във време, при което интервалът на действие или на резултата от действието съвпада с интервала на говорене, т.е. сегашно или минало неопределено време. Когато спомагателната частица *да* се употребява модално, тя често се съчетава и с минало несвършено време или минало предварително време.

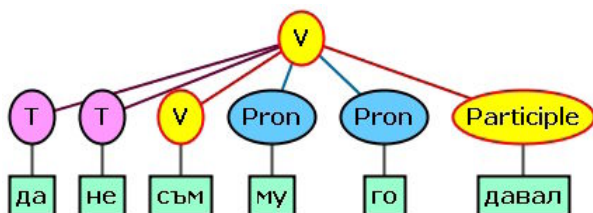
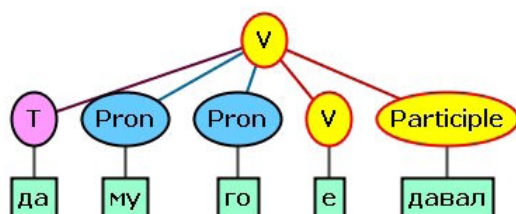
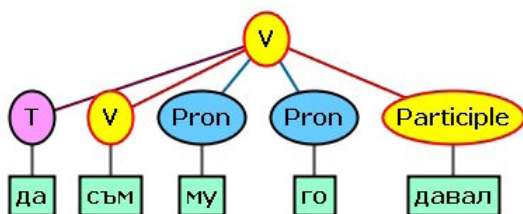
Обект на нашето внимание тук е спомагателната частица *да* в типичната си употреба. Тя избира глагол в сегашно или минало неопределено време и образува глаголен комплекс, който наследява аргументната структура на пълнозначния глагол. Самият глагол е лексикалната опора на т.нар. 'да-конструкция'. Всички клитики се реализират между *да* и глаголната форма:



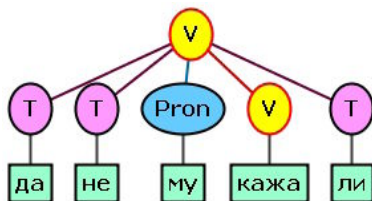
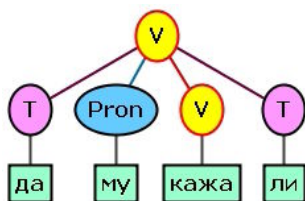
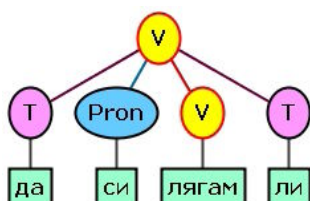
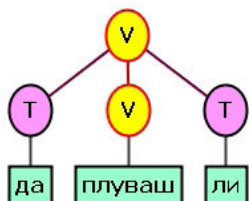
Следващите примери показват реализацията на клитиките в конструкции от *да* и по-сложни глаголни форми (в случая – минало неопределено време). Прави

¹⁵ Изключения от това правило са синонимните употреби на *да* по отношение на съюза за цел *за да* и съюза за условие *ако*: *Отивам да купя хляб* и *Да бях здрав, щях да свърша работата сам*.

впечатление подредбата да спомагателния глагол съм, който във формата си в 1 л., ед.ч. (примери 1 и 3) се реализира по-далече от глагола, докато в 3 л., ед.ч. (пример 2) се реализира непосредствено до глагола.



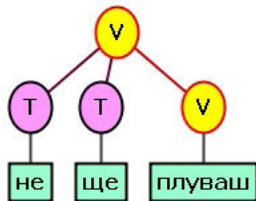
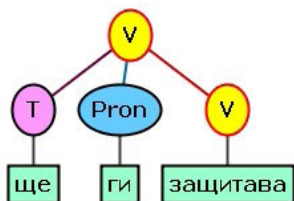
Във въпросителни изречения с частицата да в сила са същите принципи, които бяха описани по-горе. Т.е. частицата 'да-конструкцията' отляво, а след нея се реализират клитиките в подходящ словоред:



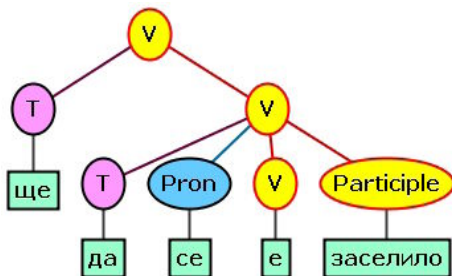
Спомагателната частица *да* в модална употреба следва модела на глаголен комплекс, който е валиден за типичната ѝ употреба. Затова не добавяме примери и за този тип конструкции.

Глаголен комплекс със спомагателната частица ‘ще’

В нашата граматика частицата *ще* също е дефинирана като ‘спомагателна глаголна частица’ (с морфологичен етикет *Tx*). Основната ѝ роля е да образува формите на бъдеще време (*ще дойда*) и бъдеще предварително време (*ще съм дошъл*). Подобно на *да*, частицата *ще* избира глагол в сегашно време или минало неопределено време. Реализацията на клитиките е същата (пример 1) с изключение на поведението на отрицателната частица *не*, която може само да рамкира отляво ‘ще-конструкцията’, а не и да влиза вътре, както е при *да* – *не да плуваш, да не плуваш, но не ще плуваш, *ще не плуваш* (пример 2). За разлика от ‘да-конструкцията’, която не образува морфологично значение от днешна гледна точка, *ще* образува. Разбира се, посоченият начин за образуване на отрицателно бъдеще време вече е архаична. В повечето случаи отрицателните форми се образуват от глагола *няма*, който избира ‘да-конструкция’ – *няма да плуваш* (вж. по-подробно в частта за глаголните фрази).



Следващият пример илюстрира друга архаизирана форма, в която участва *ще* – т.нар. презумптивни, или предположителни, форми. В този случай *ще* избира ‘да-конструкция’. Формират се две нива в рамките на лексикалния елемент. На първото ниво се образува ‘да-конструкцията’, а на следващото – презумптивната форма.



Други аналитични форми като лексикални елементи

Тук се включват всички аналитични глаголни времена с изключение на бъдеще време в миналото (*щях да ходя*) и бъдеще предварително в миналото (*щях да съм ходил*). Включват се и аналитичните наклонения, евиденциалните форми,

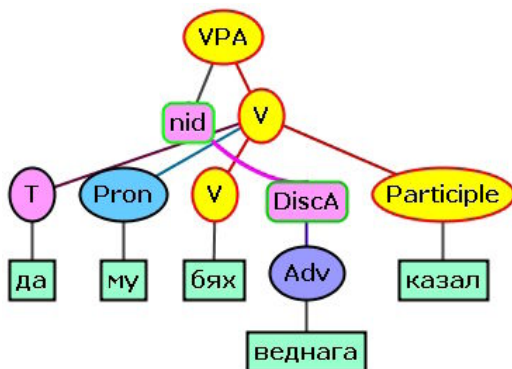
залозите. В нашата граматика *щях* е модален глагол, който избира ‘да-конструкция’ и прави глаголна фраза, т.е. не се вмества в областта на лексикалните елементи. Съзнаваме, че това разбиране противоречи на традиционното деление на категорията време на 9 грамими, но бихме искали да изтъкнем и някои допълнителни фактори. В българската граматична мисъл е имало опит да се редуцира или увеличи броят на глаголните времена според прилагането на различни критерии (Куцаров 2007, стр. 8, 204-220). От друга страна, анализът е в синхрон с анализа на да-конструкциите. От трета страна, има езикови тестове, които показват адекватността на взетото решение (по-подробно по този въпрос – вж. в раздела за глаголните фрази). Така можем да направим предположение, че категорията време се изразява не само на лексикално, но и на синтактично ниво.

Някои особености в глаголните комплекси

Освен с характера и разположението на съставлящите го елементи, глаголният комплекс е интересен и с други явления. При определени условия в него може словоредно да се появи елемент, който не принадлежи на комплекса. Също така, комплексът може да бъде пълноценен участник в кореферентна връзка с друга част на изречението (най-вече чрез кодирането на неизразен подлог в основната глаголна форма). Тези ситуации са представени тук.

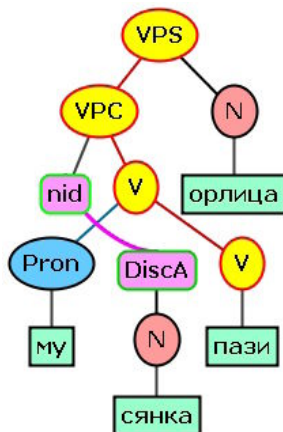
Дистантна разположеност на компонентите

Елементите на аналитичните глаголни форми могат да се разделят от други елементи. Явлението ‘дистантна разположеност’ (discontinuity) на елементи е типично за фразовото ниво, но, както виждаме, е възможно да се прояви и на лексикално ниво. Най-често мястото на прекъсване е между спомагателния глагол и причастната форма. В примера, който следва, наречието *веднага* присъства словоредно в рамките на ‘да-конструкцията’, но конституентно се намира на по-високото фразово ниво:



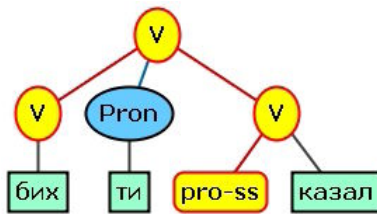
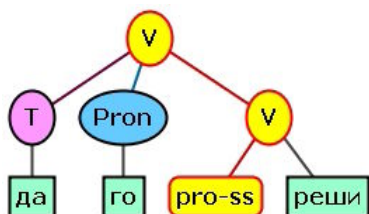
Друг, вече архаизиран случай е, когато аргумент ‘дистанцира’ клитиката от глагола. Както е в “...му сянка пази орлица”.

Разбира се, компонентите на глаголния комплекс могат да се разделят и от парентетични изрази (напр. *бях (за съжаление) заспал*).



Нулева субектност (*pro-dropness*)

Нулевата субектност е характерно явление в българския език. То се изразява в липсата на експлицитен подлог в изречението. Тук са маркирани само случаите, в които неизразеният подлог участва в кореферентна връзка. Тогава липсващата главна част е маркирана като *pro-ss* елемент. Този елемент се връзва към пълнозначния глагол или причастната форма в глаголния комплекс. Трябва да се има предвид, че той не отбелязва словоредната позиция на подлога. Примери¹⁶:



4.2

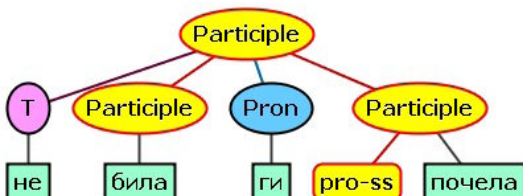
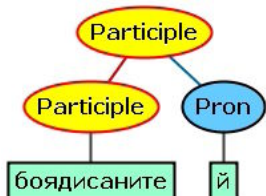
Лексикален елемент причастие

Лексикалният елемент причастие (*Participle*) съответства на самостоятелна причастна форма или на аналитичен комплекс с причастие. Самостоятелната форма е маркирана с лексикалния етикет причастие (*Participle*), ако морфологичният етикет отговаря на следния модел: “*V%%c#*”, където *V* означава глагол, *c* означава нелична глаголна форма - причастие. Представянето в графичен вид е, както следва:

¹⁶ В примерите е даден само единият член на релацията ‘кореференция’. Другият присъства в по-широкия контекст.



Тъй като причастията са хибридни елементи, те споделят характеристики както на глагола, така и на прилагателните имена. В качеството си на форми на глагола, те наследяват неговия селективен потенциал и клитики. От друга страна, поради близостта си и до прилагателното име, те свързват също притежателни клитики, които са типични за именните фрази. Някои особености са следните: отрицателната клитика става част от причастията, когато те модифицират имена (срв. *непрочетената книга*). Примери:

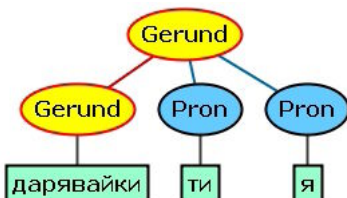


В първия пример миналото страдателно причастие пояснява име (което не е дадено в примера) и затова прозодично прилепва към себе си притежателна клитика. Вторият пример представя преизказна форма на минало неопределено или минало предварително време на глагола *почета* (*уважа*). Подреждането на клитиките е подобно на подреждането им в лексикалния елемент глагол.

4.3

Лексикален елемент деепричастие

Лексикалният елемент деепричастие (*Gerund*) съответства на самостоятелна деепричастна форма или на деепричастен комплекс с клитики. Самостоятелната форма е маркирана с лексикалния етикет *Gerund*, ако морфологичният етикет отговаря на следния модел: "**V%%g**" където **V** означава глаголна форма, **g** означава нелична глаголна форма – деепричастие. В графичен вид представянето е следното:



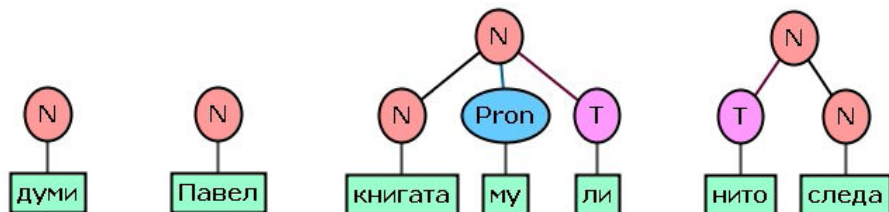
Деепричастната форма също е хибридна форма, но между глагол и наречие.

Деепричастието наследява селективната способност на глагола и затова може да привлича клитиките, които са типични за глаголните форми. Няма аналитични деепричастни комплекси. Неизразеният подлог при деепричастието (в смисъла на (Пенчев 1998, стр. 609)) е кореферентен с подлога на изречението, което модифицира¹⁷.

4.4

Лексикален елемент съществително име

Лексикалният елемент съществително име (*N*) съответства на самостоятелно съществително име или на съществително име с клитичен комплекс (*кратко притежателно местоимение, въпросителна частица, отрицателна частица*). Дадена дума се маркира като лексикален елемент *N*, ако морфологично-синтактичните характеристики отговарят на модела: “*N#*”, където **N** означава съществително име. В графичен вид представянето е, както следва:



Лексикалният елемент съществително може да бъде както нарицателно, така и собствено име (в първия и втория пример по-горе). Съществителното може да има притежателна и въпросителна клитика (трети пример) или отрицателна частица (в случая с фокусираща роля, вж. последния пример по-горе). Подобно на глаголните форми, имената могат да образуват лексикална координация, която също се смята за лексикален елемент *N*.

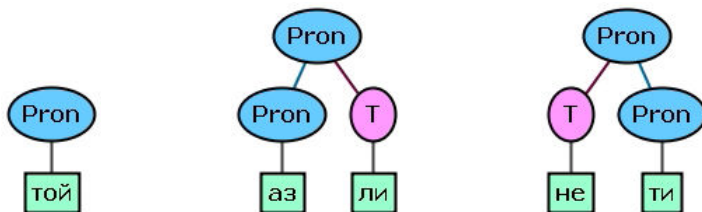
4.5

Лексикален елемент местоимение

Лексикалният елемент местоимение (*Pron*) съответства на самостоятелно местоимение или на местоимение, придружено от клитики (*кратко притежателно местоимение, въпросителна частица, отрицателна частица*). Известно е, че терминът ‘местоимение’ не е точен, тъй като членовете на тази част на речта се употребяват и вместо наречия, дори глаголи. Затова, макар да използваме традиционния термин, всъщност имаме предвид ‘местоименни

¹⁷ В някои случаи се наблюдава кореференция на деепричастието с непряк обект, когато семантичната му роля е експериенцер. Напр. *Гледайки през прозореца, на него му се стори, че пада*. (Този аспект от поведението на деепричастието ни беше посочен от проф. Ницолова в личен разговор.)

думи'. При моделирането им се придържахме към концепцията на (Ницолова 1986). Разделяме ги на три основни групи: *субстантивни*, *адективни* и *адвербиални местоимения*. Тук представяме само първата група. Другите две са представени съответно в частите за прилагателното и наречието¹⁸. Една дума се маркира с лексикалния етикет местоимение (*Pron*), ако отговаря на следните модели: "**Pp#**", "**Pde#**", "**Pre#**", "**Pce#**", "**Pie#**", "**Pfe#**", "**Pne#**", "**Psot#**", "**Pszt#**", "**Psxt#**" където **P** означава местоимение, **p** означава лично местоимение, **d** означава показателно местоимение, **r** означава относително местоимение, **c** означава обобщително местоимение, **i** означава въпросително местоимение, **f** означава неопределително местоимение, **n** означава отрицателно местоимение, **s** означава притежателно местоимение, **e** означава същини (лица и други обекти), **o** означава един притежател при притежателните местоимения, **x** означава възвратност, **z** означава много притежатели при притежателните местоимения, **t** означава кратка форма. В графичен вид представянето е следното:



Първият пример илюстрира най-простия случай, в който едно местоимение е маркирано с лексикален етикет. Вторият пример показва последователността на лично местоимение и въпросителна клитика. Третият пример съответно показва отрицателна клитика към местоимението.

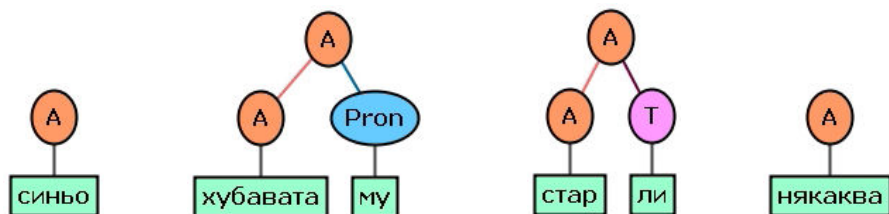
4.6

Лексикален елемент прилагателно

Лексикалният елемент прилагателно (*A*) съответства на самостоятелно прилагателно или прилагателно с прилежащи клитики, които са същите като за другите именни елементи. Една дума се маркира с лексикалния етикет прилагателно (*A*), ако отговаря на следните модели: "**A#**", "**Pda#**", "**Pra#**", "**Prp#**", "**Pca#**", "**Pcq#**", "**Pia#**", "**Pip#**", "**Pfa#**", "**Pfp#**", "**Pna#**", "**Pnp#**", "**Psol#**", "**Pszl#**", "**Psxl#**", където **A** означава прилагателно, **P** означава местоимение, **a** означава качество, **c** означава обобщително местоимение, **d** означава показателно местоимение, **f** означава неопределително местоимение, **i** означава въпросително местоимение, **l** означава дълга форма, **n** означава отрицателно мес-

¹⁸ Трябва да обърнем внимание, че морфологичните етикети, изброени в обобщен вид тук, описват всички видове местоимения и отразяват произхода им като част на речта. Лексикалният етикет е абстракция на тяхната функция. Место-съществителните се маркират с лексикален етикет *Pron*, место-прилагателните – с *A*, а место-наречията – с *Adv*. По този начин запазваме информацията, че думата е местоимение със своите характеристики като част на речта, но в същото време има определена синтактична функция.

тоимение, **o** означава един притежател при притежателните местоимения, **p** означава притежателност, **q** означава количество, **r** означава относително местоимение, **s** означава притежателно местоимение, **x** означава възвратност, **z** означава много притежатели при притежателните местоимения. В графичен вид представянето е следното:

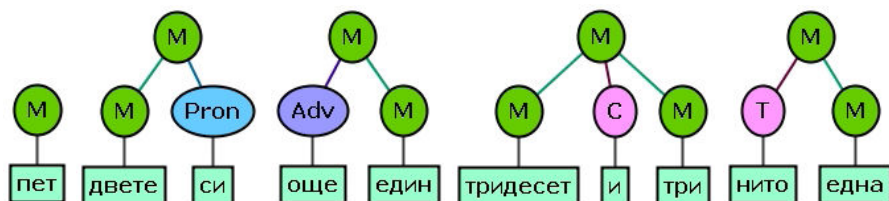


Първият пример представя лексикалната проекция на самостоятелно употребено прилагателно. Вторият пример показва комбинация от членувано прилагателно и прилежаща клитика – кратко притежателно местоимение. Третият пример демонстрира комбинация от прилагателно с въпросителна клитика. Последният пример показва лексикална проекция на адективно местоимение от реда на неопределителните местоимения.

4.7

Лексикален елемент числително

Лексикалният елемент числително (**M**) съответства на самостоятелно числително, аналитични комбинации от числителни или числителни от двата вида с прилежащи клитики. Аналитично изразените числителни се състоят от числителни, свързани със съюза 'и' (напр. *тридесет и осмият*). Една дума се маркира с лексикалния етикет **M**, ако отговаря на следния модел: "**M#**", където **M** означава числително име. В графичен вид представянето е следното:

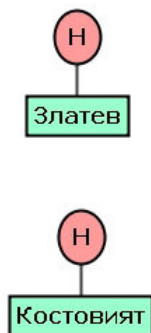


Първият пример представя по правило базовия случай, при който едно числително има лексикална проекция на числително име. Вторият пример представя членувано числително с възвратнопритежателна клитика. Третият пример показва числително име с фокусираща дума. Четвъртият пример показва моделирането на аналитично числително име. Петият пример показва комбинацията от числително име и отрицателна фокусираща дума. Разбира се, числителните могат да бъдат представени и чрез цифри: 30, 3.14159, 2001, 1994-1996, 1,72, 30 000, 18:23; 7-и; XII.

4.8

Лексикален елемент фамилно име или прилагателно, образувано от собствени имена

Този лексикален елемент няма съответствие сред традиционните части на речта. Той беше въведен поради следните причини: необходимостта да се анализират подобни примери по-детайлно, както и поради генеалогичната връзка между прилагателните от собствени имена и фамилните имена. Затова лексикалният елемент (*H*) съответства на самостоятелно фамилно име или прилагателно от собствено име. Разбира се, тук се включват и презимената, тъй като те следват същия словообразователен тип. Една дума се маркира с лексикалния етикет *H*, ако отговаря на следния модел: "**H#**", където **H** означава фамилно име или прилагателно от собствено име. В графичен вид представянето е следното:

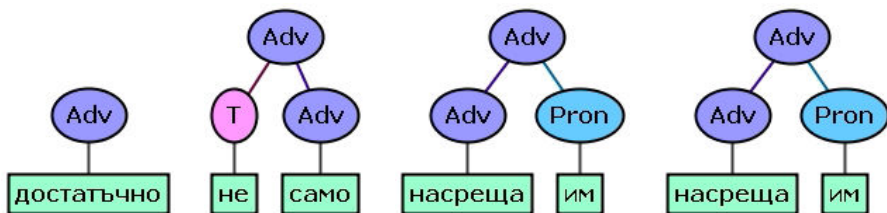


Първият пример демонстрира лексикалния елемент *H*, който може да бъде фамилно име или прилагателно от име – формата е двузвучна и зависи от контекста на употреба (напр. *г-н Златев*, но: *Златев съвет*). Вторият пример показва прилагателно, образувано от лично име.

4.9

Лексикален елемент наречие

Лексикалният елемент наречие (*Adv*) съответства на самостоятелно наречие, самостоятелно местоименно наречие или на наречие с клитики (въпросителна, отрицателна, дателна). Една дума се маркира с лексикалния етикет *Adv*, ако отговаря на следния модел: "**D#**", "**Pd%**", "**Pr%**", "**Pc%**", "**Pi%**", "**Pf%**", "**Pfq#**", "**Pfy%**", "**Pn%**", където **D** означава наречие, **P** означава местоимение, **c** означава обобщително местоимение, **d** означава показателно местоимение, **f** означава неопределително местоимение, **i** означава въпросително местоимение, **n** означава отрицателно местоимение, **q** означава количество, **r** означава относително местоимение, **y** означава количество за лица. В графичен вид представянето е следното:

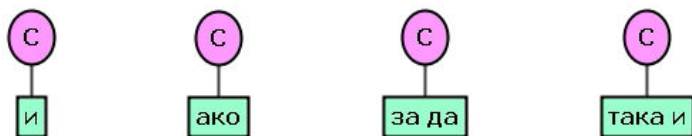


Първият пример представя самостоятелно употребено наречие. Вторият пример показва наречие с отрицателна частица. Третият пример е илюстрация за местоименно наречие с въпросителна клитика. Последният пример показва наречие с дателна клитика. Съществуват сложни наречия, които се анализират като многокомпонентни лексикални елементи. Такива са: *така нататък, все пак, все едно, може би, едва ли, едва ли не*.

4.10

Лексикален елемент съюз

Лексикалният елемент съюз (С) съответства на частта на речта - съюз. Една дума се маркира с лексикалния етикет С, ако отговаря на следния модел: "С#", където С означава съюз. Съществуват и сложни съюзи, които се анализират като многокомпонентни лексикални елементи. В графичен вид представянето е следното:



Горните илюстрации показват делението на съюзите на групи по два критерия: прости (примери 1 и 2) и сложни (примери 3 и 4); съчинителни (примери 1 и 4) и подчинителни (примери 2 и 3).

Представяме и списък с думи, които според нас са в групата на сложните съюзи: *ето защо, за да, както и, камо ли, освен ако, освен като, освен че, само ако, само че, след като, сякаш че, тъй като, тъй че, така и, така че, като че, като че ли*¹⁹.

Трябва да отбележим, че нашият списък е редуциран в сравнение с традиционното схващане за сложен съюз. По принцип сложните съюзи се образуват от предлог + съюз (*освен ако, след като*), наречие + съюз (*тъй като, сякаш че, както и*), други модели (*ето защо, за да, камо ли*). Нашият списък е получен чрез прилагане на тест за вмъкване на лексикален елемент между частите на кандидатите за сложни съюзи. Ако може да се вмъкне елемент, тогава смятаме последователността от думи не за комплексна цялост, а за независими елементи. Ако обаче не може да се вмъкне елемент, със сигурност става дума за сложен съюз. Напр. *преди да* не е сложен съюз, а предлог + да-конструкция, тъй като може да се вмъкне дума между двете части: *Преди Иван да дойде...*

¹⁹ Определянето на лексикалния етикет на елементите: *като че* и *като че ли* е в зависимост от контекста. В изречението *Той се засили, като че ли щеше да хвъркне думата като че ли* е съюз. Но в изречението *Той като че ли страдаше за случилото се думата като че ли* е наречие.

От друга страна, *за да* е сложен съюз, тъй като не може да се вмъква друг лексикален материал между частите му: **за Иван да дойде*. Разбира се, има хибридни случаи, при които в определени условия може да се вмъкне лексикален материал, но не е типично да се вмъква. Такива са напр. съюзите *освен ако*, *освен като*, *само ако* от горния списък.

4.11

Лексикален елемент предлог

Лексикалният елемент предлог (*Prep*) съответства на частта на речта - предлог. Една дума се маркира като *Prep*, ако морфологично-синтактичните ѝ характеристики отговарят на етикета: “R”, където R означава предлог. Съществуват сложни предлози, които се анализират като многокомпонентни думи. В графичен вид представянето е следното:



Първите два примера представят прости предлози, докато вторите два показват сложни. Сложните предлози обикновено имат следната структура: предлог + съществително име (*по случай*), предлог + съществително име + предлог (*в зависимост от*), съществително + предлог (*благодарение на*).

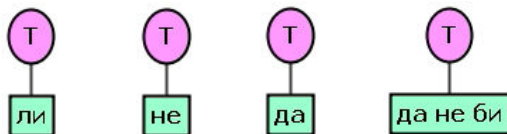
Ето списък на някои по-чести последователности, които в една или друга степен са кандидати за етикета ‘сложни предлози’: *във връзка с*, *в сравнение с*, *в зависимост от*, *в съгласие с*, *в хода на*, *с оглед на*, *по отношение на*, *в съответствие с*, *в близост до*, *за сметка на*, *по време на*, *в полза на*, *в интерес на*, *по случай*, *в резултат на*, *от страна на*, *по линия на*, *в памет на*, *благодарение на* и др. Трябва да се вземе предвид фактът, че макар да наричаме гореизброените предлози сложни поради тенденцията им към идиоматизиране, между компонентите има различна степен на споеност. Нека например приложим тест, при който съчетанието име + предлог се заменя с притежателно местоимение. Сложните предлози *от страна на*, *в памет на*, *за сметка на*, *в интерес на* и др. могат да се парафразират като *от негова страна*, *в негова памет*, *за тяхна сметка*, *в мой интерес*, което показва по-голяма споеност между съществителното и следващия го предлог. Те позволяват композиционалност при анализа: [PP в [NPA памет [PP на поета]]]. Подобен анализ обаче не можем да приложим в следния случай: *[PP с [NPA оглед [PP на причините]]], което означава, че съчетанието *с оглед на* (=предвид на) е достигнало по-висока степен на абстрактност и е по-добър кандидат за сложен предлог. Има много други фактори, които влияят върху типологията на сложните предлози – напр. морфологично-синтактичните особености и поведение на съществителното, възможността да се вмъква модифициращ елемент и др.

Въпросът се нуждае от по-задълбочено проучване.

4.12

Лексикален елемент частица

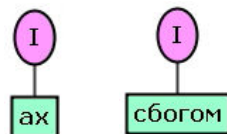
Лексикалният елемент частица (*T*) съответства на частта на речта – частица. Една дума се маркира като (*T*), ако следва модела: “*T#*”, където *T* означава частица. Всъщност следващата буква посочва типа на частицата (утвърдителна, спомагателна и др. – вж. приложението). Срещат се и сложни частици, напр. *да не би*, която се анализира като многокомпонентна дума. В графичен вид представянето е следното:



4.13

Лексикален елемент междуметие

Лексикалният елемент междуметие (*I*) съответства на частта на речта – междуметие. Една дума се маркира с етикета (*I*), ако е записана като междуметие в лексикона. В графичен вид представянето е следното:



И двата примера представят типични междуметия: първи пример – емоция, а втори пример – формула при сбогуване, разделяне. Но в този лексикален елемент се включват и звукоподражателните думи като *мяу*, *туп*, *скръц* и под.

4.14

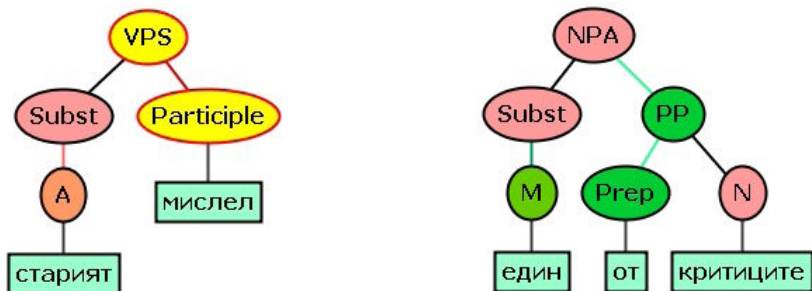
Промяна на типа (type shifting)

Промяната на типа може да се дефинира най-общо по следния начин: семантична операция, която превръща аргумент в типа, който се очаква от функцията, тъй като иначе ще се получи грешка²⁰ (Pustejovsky 1995, p. 59). За нашите цели използваме модифициран вариант на горното определение, който гласи: смяна на типа наричаме явление, което означава, че една дума или фраза от дадена категория се използва като лексикален елемент на друга категория. Тук разглеждаме 3 основни явления: *субстантивация*, *номинализация* и *вербализация*.

²⁰ Преводът от английски е мой – П. О.

Субстантивация

Това явление се основава предимно на контекста и представя промяна на зависимите части от именната фраза в опори, които обикновено се изразяват чрез съществителни имена. Характерно е, че това явление винаги е лицензирано от елипса на опората, която принуждава зависимата част да поеме ролята на тази опора. На субстантивация се подлагат именни елементи като прилагателни, числителни, причастия, местоимения. В графичното представяне подобни елементи са маркирани като *Subst*. Примери:



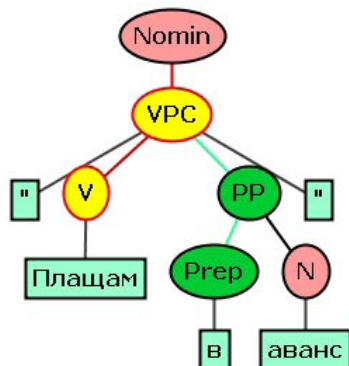
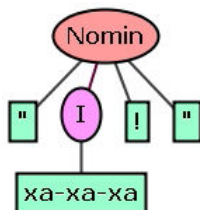
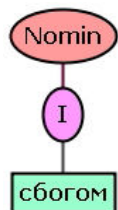
Първият пример показва субстантивация на прилагателно. При него липсва опората, която обаче може да се възстановява в контекста (напр. *старият лесничей*) или по-общо в рамките на предикативната структура (напр. от семантиката на глагола *мисля* е ясно, че липсващата опора реферира към лице – *старият човек*). Вторият пример е дискуссионен и по-сложен за анализ. Има две възможни лингвистични решения: *един* да се субстантивира или да се маркира елипса на *критик* (*един критик от критиците*). Ние избрахме първата възможност не защото лингвистично е по-добра от другата, а защото елипсата е възстановима в много тесен контекст и няма опасност от загуба на информация. Освен това, от чисто имплементационна гледна точка се спестява въвеждането на допълнителни елементи.

Номинализация²¹

Подобно на субстантивацията, номинализацията също се лицензира от елипса на опората, но елипса на модификатор. За разлика от субстантивацията, номинализацията променя основно предикати, междуметия и други не-именни фрази в именни. Номинализацията се маркира като елемент *Nomin*.

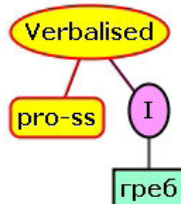
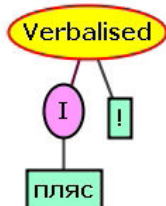
²¹ Трябва да обърнем внимание, че тук не разглеждаме анафорично-катафоричната номинализация, при която местоименен елемент отправя и номинализира фраза или предикат. Напр. *Той влезе и седна на пода – това ме ядоса*. Показателното местоимение *това* номинализира цялата предикатна структура *той влезе и седна*.

Следват примери за номинализация на междуметие (първи и втори пример) и на предикат, който е название на пиеса (трети пример):



Вербализация

Този подвид промяна на типа се проявява, когато междуметия, частици или наречия изпълняват в контекста функция на предикати. Най-честите случаи са, когато междуметия влизат в тази роля. В синтаксиса те се наричат *междуметни сказуеми*. Елементът, подложен на вербализация, се маркира с етикета *Verbalized*. Примери:

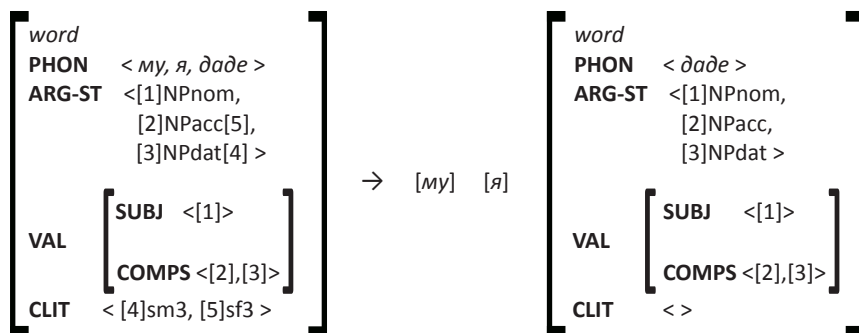


Първият пример представя основния случай. Вторият пример показва вербализиран елемент с маркер за неизразен подлог, което означава, че елементът наистина е поел функциите на глагола и че в тази си функция може да участва в механизмите на кореференция. За пълнота трябва да споменем, че тези вербализирани елементи могат да имат и свои компоненти. Напр. *И той лап! кюфтенце!*

Глаголна фраза (вербална фраза)

Глаголната област се състои от лексикални елементи *V* и фразови елементи: опора-комплемент – *VPC(omplement)*, опора-подлог – *VPS(subject)*, опора-адюнкт – *VPA(djunct)* и опора-запълваща част – *VPF(iller)*. В този раздел представяме фразовите елементи от глаголната област. Следвайки общата йерархия на реализация на зависимите части, се придържаме към следната поредност: *лексикален глагол* → *глаголна фраза от тип опора-комплемент* → *глаголна фраза от тип опора-подлог* → *глаголна фраза от тип опора-адюнкт*. Тук те ще бъдат разгледани в тази подредба. За лексикалните глаголни форми вижте предишния раздел.

Както стана ясно от раздела за Опорната фразова граматика, при глаголните фрази е много важно взаимодействието между аргументния списък и валентните списъци. Фразите от тип *VPC* и *VPS* се управляват от Валентния принцип до пълното реализиране на аргументите на опората. Фразата от тип *VPA* се управлява от Принципа на адюнкта, където самият адюнкт избира опората си чрез **MOD** характеристиката. Фразата от тип *VPF* се управлява от Принципа на запълващата фраза. Характерното за българския език дублиране на аргументите изисква по-специално внимание към реализацията на клитиките и връзката им с изразените компленти. Ето пример с изречението *Иван да* [_i му] [_j я] *даде* [_i на Петър] [_j картината]. Както казахме в предишния раздел, клитиките според нас се реализират на лексикално ниво. При тази своя реализация те модифицират характеристиките на опората. Дателните и винителните клитики си взаимодействат с аргументната структура на глагола и образуват нова аргументна структура, в която характеристиките на клитиките са отразени в рамките на съответните аргументи²². Тук даваме правилото за реализация на тези две клитики.



²² Клитиките, които не са свързани с аргументи, модифицират други характеристики на опората и чрез други механизми.

Правилото показва само схематично ролята на съответните клитики. Първо, те се реализират като елементи на списъка, който е стойност на атрибута **CLIT**. Оттам информацията, която се носи от клитиките, се разпространява на две места – като фонетичен материал в рамките на стойността на атрибута **PHON** и като граматични характеристики към съответните аргументи в рамките на аргументната структура на новия знак. Това става чрез индексите [4] и [5].

Следва разглеждане на различните типове глаголни фрази според вида на зависимите части.

5.1

Глаголна фраза от типа опора-комплемент – VPC

Всяка фраза от тип *VPC* има опора и една или повече комплементни дъщери. Всички комплементи се реализират наведнъж. Следните лексикални елементи могат да бъдат опори: *V*, *Participle*, *V-Elip*, *VD-Elip*, *CoordP*²³ и *Verbalised*. Лексикалните елементи *V*, *Participle* и *Verbalised* са описани в предишния раздел. Да споменем, че координацията е лексикална, когато нейните аргументи са лексикални (вж. раздела, посветен на лексикалните елементи по-горе, и раздела, посветен на координацията по-нататък). Елементите *V-elip*, *VD-elip* представляват глаголни елипси в изречението и са обяснени в раздела, посветен на елипсите. Те са лексикални, когато липсващият елемент е лексикален. Комплементите включват следните типове фрази: *именни елементи*, *адвербиални елементи*, *причастия*, *предложни фрази*, *подчинени изречения*, *координационни фрази*. По-нататък дискутираме различните случаи. Опората може винаги да бъде определена еднозначно, тъй като нашето множество от категории е дефинирано така, че множеството на опорите и множеството на комплементите не се пресичат²⁴. Когато всички дъщери на една *VPC* фраза са в координационни отношения, то опората и комплементите се определят на базата на категориите на аргументите на съответните координации.

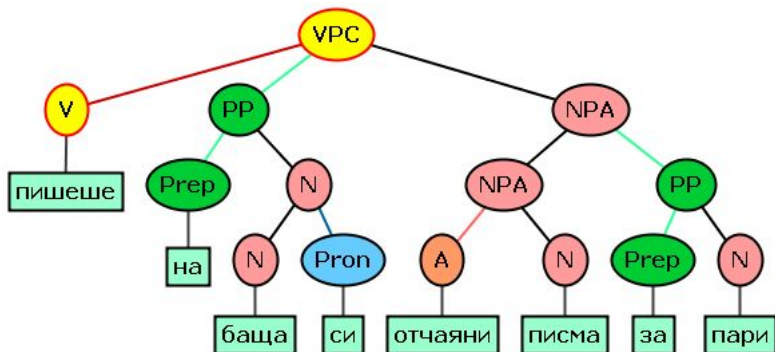
Комплементите се реализират в произволна словоредна подредба спрямо опората²⁵. Някои комплементи могат да бъдат реализирани словоредно извън обсега на *VPC* фразата и те се маркират локално като *nid* елемент. Ако лек-

²³ Фразата *CoordP* се отнася както за фразовата, така и за лексикалната координация. Затова видът на координацията се разбира не от етикета, а от вида на аргументите ѝ.

²⁴ Прави впечатление, че причастието присъства в множеството на опорите и в множеството на комплементите. Пресичане обаче няма поради следните причини: първо, сегашните деятелни причастия (*той е пеещ*) не участват в образуването на глаголни форми, следователно не могат да бъдат опори и второ, другите видове причастия, които се появяват в тази позиция, се възприемат по-скоро като омоними-прилагателни. Срв. *Детето е разглезено от родителите си.* с *Детето е разглезено. (=капризно).* Всякакви други последователности от причастия не са многозначни по отношение на функцията си. Морфологичният етикет дава информация кое причастие идва от спомагателен и кое – от пълнозначен глагол.

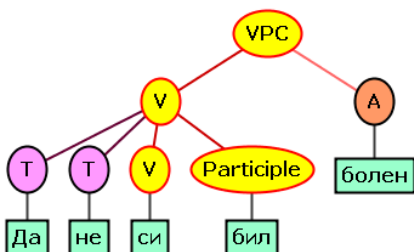
²⁵ Разбира се, произволна в рамките на ограниченията от граматиката на българския език.

сикалната опора съдържа дателни или винителни клитики, то реализираните компленти трябва да споделят с тях стойностите на характеристиките число, род и лице. Опората и комплементите могат да се реализират дистантно при определени условия. Те могат да бъдат разделени от *DiscA* елементи, които може да са подлогът или адюнктите на същата лексикална опора, или *Pragmatic* елементи като парентетични изрази, звателни форми и други. Броят и видът на комплементите се определя от лексикалната опора. Следват няколко по-обща примера, а след това са представени и някои по-специфични случаи:



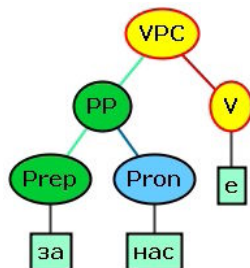
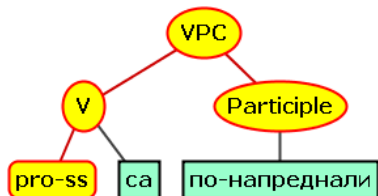
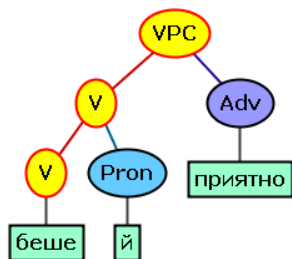
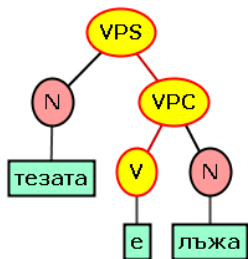
В този пример има глаголна фраза от тип *опора-комплент* с лексикална опора *пишеше*, която изисква два комплента – дателен непряк обект (*пиша на някого*), изразен от предложната фраза *PP* – и винителен пряк обект (*пиша нещо*), изразен от именна фраза *NPA*.

В следващия пример опората е сложен глаголен комплекс, представляващ ‘да-конструкция’. Тя включва минало неопределено време на глагола *съм*, глаголната спомагателна частица *да* и отрицателната частица *не*. Комплементът в този случай е прилагателно.



Глаголна фраза опора-комплент с опора спомагателен глагол ‘съм’ (копула)

Приемаме идеята, че всички предикативни аргументи (именни фрази, предложни фрази и адвербиални фрази) на копулата са нейни компленти. Приемаме също така за дадено, че комплементът на копулата се кореферира с подлога ѝ в случаите на предикативно съгласуване. Напр. *Стаята е хубава*, но не: **Стаята е хубав*. Това явление е подобно на случаите на вторична предикация, където подлогът или комплементът се съгласува с адюнкта, участващ във вторичната предикация. Тук показваме някои от най-честите типове компленти при копулата:



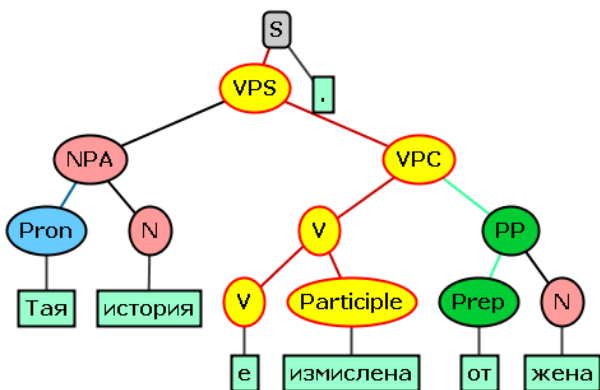
Тези примери показват различни типове компленти на копулата – номинален елемент, наречие, причастие в предикативна употреба и предложна фраза. Примерите показват и различни форми на опората. Тя може да бъде свързана с клитика или да има елемент *pro-ss* в случаите, когато подлогът е неизразен, но участва в кореферентно отношение в рамките на изречението.

Като фрази от тип *опора-комплент* се разглеждат и предикативи с наречия като *ясно е*, *добре е*, *интересно е* и подобни. Понякога копулата може да бъде изпусната. Напр. изречението с липсваща копула *Естествено, че ще дойда* е твърдествено на изречението с присъстваща копула *Естествено е, че ще дойда*. Тези примери са разгледани подробно по-нататък в раздела, посветен на явлениято *елипса*.

Конструкции със страдателен залог при преходни глаголи

При конструкциите със страдателен залог винаги разглеждаме логическия подлог, изразен чрез предложна фраза (най-често с предлог *от*), като комплент. Промяната на позицията на подлога от деятелна форма в позиция на комплент при страдателната форма се осъществява с помощта на лексикално правило.

Със съответните синтактични модификации то разменя местата на фразата, изпълняваща граматичната роля на пряко допълнение (винителния комплент) и на фразата, изпълняваща ролята на подлог. Следва пример на страдателния залог на изречението: *Жена измисли тая история*.

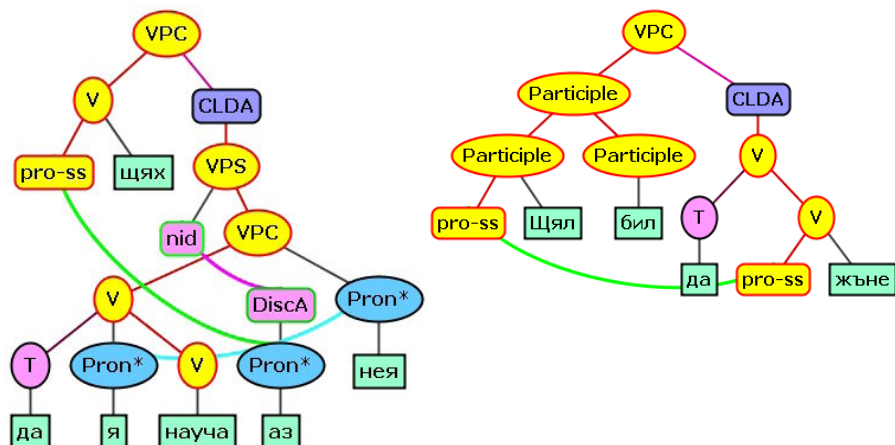


Аналитични глаголни форми като VPC

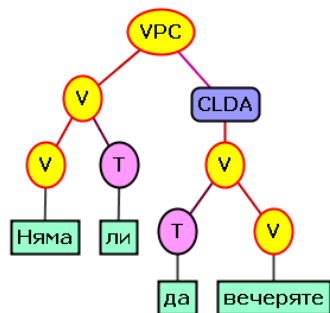
Както беше споменато в раздела за лексикалните глаголи, някои традиционни аналитични глаголни форми са разгледани при нас като глаголи, които изискват подчинени изречения за компленти. Тоест когато пълнозначният глагол е в 'да-конструкция', приемаме, че 'да-конструкцията' е опора на подчинено изречение (с изразени или не подлог, компленти и адюнкти) и това подчинено 'да-изречение' се изисква като комплемент от спомагателния глагол. Такива форми са аналитичните глаголни форми със спомагателния глагол *ща* в минало време – *щях* (бъдеще в миналото, бъдеще предварително в миналото и техните преизказни форми) и отрицателните форми за бъдеще време – *няма*, *нямаше* и *нямало*²⁶. С други думи, това са формите, след които се появява 'да-конструкция'. Мотивацията за това решение е, че в случаите на изразени подлог и други зависими части на 'да-конструкцията', те се реализират локално и образуват подчинено 'да-изречение'. Това подчинено 'да-изречение' може бъде реализирано и в други контексти, без да промени своята синтактична структура и семантика. Спомагателните изречения от този тип се маркират с функционалния елемент *CLDA*. В първия пример по-долу подлогът и комплементът са реализирани в рамките на 'да-изречението'. Тъй като според йерархията за реализация на зависимите части първо се реализира комплементът, подлогът *аз* разделя опората и комплемента. По тази причина подлогът е маркиран с функционалния елемент *Disca* и неговото участие в конституентната структура на по-високата фраза е указан чрез елемента *nid*. Обърнете внимание на двете кореферентни релации: първо, между комплемента и винителната клитика в рамките на 'да-изречението' и второ, между изразения подлог на 'да-изречението' и неизразения подлог на спомагателния глагол *щях*, който е представен с помощта на еле-

²⁶ Осъзнаваме, че на пръв поглед се получава абсурдната ситуация бъдеще време в положителната си форма да е лексикален елемент (*ще дойда*), а в отрицателната си да е фразов (*няма да дойда*), но смятаме, че: (1) има лексикален елемент за отрицателната форма, който обаче вече не е толкова често употребяван (*не ще дойда*); (2) езикът предлага различни варианти за описание на езиковите обекти. Ако приемем, че глаголните времена в българския език демонстрират както лексикална, така и фразова структура, граматиката остава непротиворечива.

мента *pro-ss*. Вторият пример показва кореференция между двата неизразени подлога на спомагателния глагол *щял* и на подчиненото 'да-изречение'.



Последният пример показва отрицателна форма за бъдеще време. Примерът представя също и прозодичната комбинация на въпросителната частица *ли* и отрицателния глагол като лексикален глаголен знак. В този пример *няма* е в безлична употреба и поради тази причина няма кореферентна релация с неизразения подлог на 'да-изречението'.

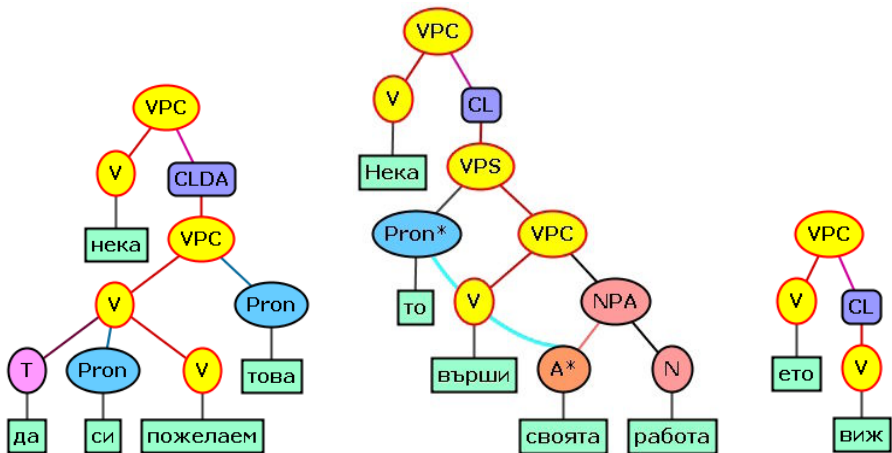


Подчинени изречения като компленти

Тук се разглеждат различните по тип подчинени изречения. От тях като по-чести сме отделили подчинени изречения с 'да-конструкция', маркирани с функционалния етикет *CLDA*; подчинени изречения, въведени със съюза *че* – *CLCHE*; подчинени изречения, въведени със съюза *за да* или *да* със същото значение – *CLZADA*; въпросителни подчинени изречения – *CLQ*; относителни подчинени изречения – *CLR* и други подчинени изречения, маркирани с общ функционален етикет – *CL*. Ролята на всеки вид подчинено изречение като комплемент се дискутира по-долу в раздела, посветен на подчинените изречения. Затова тук коментираме само някои случаи, при които опората е глаголна частица или глагол с по-особено поведение.

Глаголните частици като *нека* и *ето*²⁷ взимат като компленти подчинени изречения от тип *CLDA* или *CL*. Ние маркираме тези частици като лексикални глаголни знакове.

²⁷ Частицата *ето* може да взима и други типове компленти.



В първия пример частицата *нека* избира подчинено ‘да-изречение’ като комплемент. Във втория пример тя избира общ вид подчинено изречение без ‘да-конструкция’. В последния пример частицата *ето* избира също подчинено изречение от общ тип.

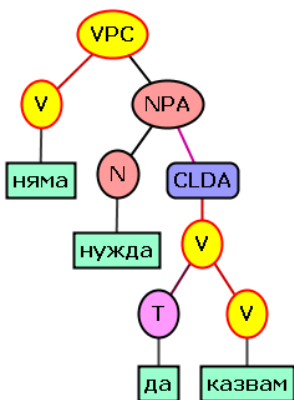
Към глаголите с по-особено поведение отнасяме глагола *стига*, който може да избира като комплемент подчинени изречения от общ тип (напр. *стига си спал*). При комбинация с подчинено ‘да-изречение’ цялата конструкция става условна и може да се приеме, че *стига* играе ролята на съюза *ако*. В този случай *стига* въвежда условно подчинено изречение, което е в позицията на адюнкт спрямо главното. Напр. **Стига да** искам, *ще успея*, което е синонимно на изречението **Ако** искам, *ще успея*.

Специални случаи

Това са конструкции с т.нар. *поддържащи* (support) или *леки* (light) глаголи.

Тези глаголи следват модела *глагол+име* и заедно правят комплексна предикация, т.е. глаголът е десемантизиран до известна степен и затова се нуждае от семантиката на името (*правя впечатление* = *впечатлявам*), но все пак не е загубил напълно способността си да предичира (*хвърлям поглед* = *поглеждам*, но *хвърлям топа* = *умирам*). Други примери: *държа реч пред студентите*, *държа врага под око*.

Глаголите *имам*, *нямам* и подобни на тях в употреба на поддържащи глаголи формират VPC от специален вид. Вместо традиционния анализ, при



който те взимат два комплемента, ние предлагаме друг анализ. При него, техният първи комплемент взема втория като свой модификатор. Обикновено първият аргумент е именна фраза, която е модифицирана от втория аргумент. В примера съществителното се свързва с ‘да-изречението’ и образува фраза *NPA*, която вече в своята цялост става комплемент на глагола.

5.2

Глаголна фраза от типа опора-подлог – *VPS*

Фразите от този тип имат опора и един подлог. *VPS* може да има лексикална опора, ако тази лексикална опора не изисква комплемент. Следните лексикални и фразови елементи могат да бъдат опори: *VPF*, *VPC*, *V*, *Participle*, *V-Elip*, *VD-Elip*, *CoordP* и *Verbalised*, тоест всички възможни опори за *VPC* са възможни опори и за *VPS*, а също така опора може да бъде и самата *VPC*. Фраза от тип *VPF* може да бъде опора на *VPS* фраза, ако запълващата фраза се реализира пониско от подлога. Лексикалните елементи *V*, *Participle* и *Verbalised* са описани в раздела за лексикалните елементи по-горе.

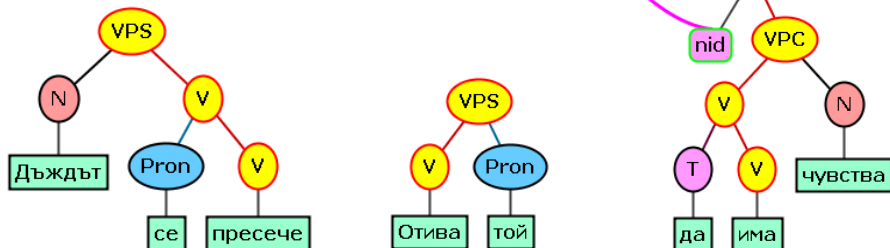
Подлогът може да бъде *именна фраза*, *подчинено изречение*, *предложна фраза*, *адвербиална фраза* или *координационна фраза*. Според нас някои предложни и адвербиални фрази също имат роля на подлог, когато изразяват приблизително количество на обекти или материали²⁸. Напр. *Около 50 студенти слушаха лекцията*. В тези случаи приемаме, че те имат номинален референт. Подобно на случая с глаголните фрази от тип опора-комплемент, и тук опората може винаги да бъде определена еднозначно, тъй като нашето множество от категории е дефинирано така, че множеството на опорите и множеството на възможните подлози не се пресичат. Когато всички дъщери на една *VPS* фраза са координационни, то опората и подлозите се определят на базата на категориите на аргументите на съответните координации.

Подлогът се реализира в произволна словоредна подредба спрямо опората. Разбира се, има локални ограничения, като например: той не може да бъде в рамките на ‘да-конструкцията’ на глагола (**да Петър дойде*). Подлогът може да бъде реализиран словоредно извън обсега на *VPS* фраза и тогава се маркира локално като *nid* елемент. Подлогът обикновено се съгласува с глаголната опора по число, род и лице²⁹. Опората и подлогът могат да се реализират дистантно. Те могат да бъдат разделени от безличен глагол, *DiscA* елементи, които може да са адюнкти на същата лексикална опора; *DiscE* елементи, които може да са конституенти на комплементите на същата лексикална опора, или *Pragmatic* еле-

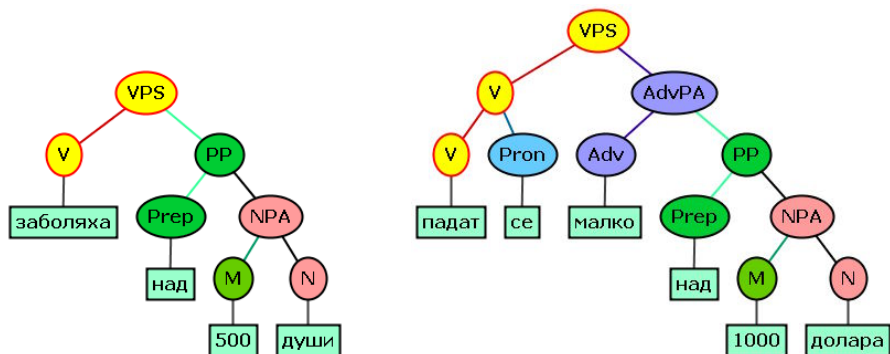
²⁸ Другата алтернатива за анализ на подобни конструкции е да се разглежда последователността предлог+числително *около 50* като сложен спецификатор към опората *студенти*. Тогава се избягва неудобството предложна фраза да се разглежда като подлог. И двете възможности обаче, доколкото ни е известно, имат своите предимства и недостатъци.

²⁹ За изключения от това правило вж. (Граматика на съвременния български книжовен език. Синтаксис 1983, стр. 144-154).

менти като парентетични изрази, звателни форми и други. Самият подлог може да се реализира като *DiscA* елемент между опората и нейните компоненти. Видът на подлога се определя от лексикалната опора чрез нейната аргументна структура (атрибут **ARG-ST**). Следват примери на някои реализации на подлога:



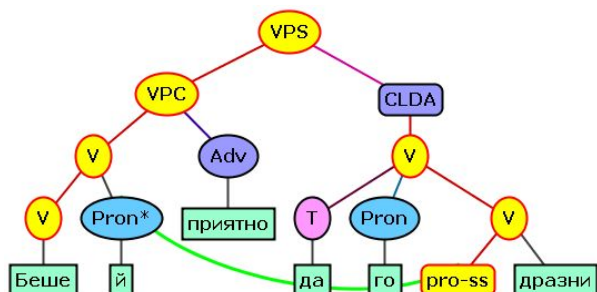
Първият пример показва съществително като подлог. Словоредът е обективен, т.е. изречението следва линейния модел подлог-сказуемо. Вторият пример показва местоименен подлог, който словоредно се появява след опората. Третиият пример демонстрира реализацията на т.нар. 'извлечен подлог' в конструкция с безличния глагол *трябва*, който избира подчинено 'да-изречение'. Механизмът на извличане се налага поради факта, че безличният глагол няма подлог. Следователно съществителното *човек* е подлог на 'да-конструкцията' *да има чувства*, а не на глагола *трябва*. Затова то е маркирано като *DiscE* елемент и е доминирано от *VPF* фразата.



Горните два примера показват подлог, изразен с предложна фраза и с адвербиална фраза. Ако игнорираме предлога и наречието, оставащата именна фраза се съгласува с опората.

Въпросът за подчиненото изречение в ролята на подлог се разглежда по-нататък в частта за подчинените изречения. Тук за пълнота даваме един пример

на подчинено подложно изречение. Обърнете внимание, че неизразеният подлог на подчиненото ‘да-изречение’ е в кореферентна връзка с дателната клитика в рамките на опората на главното изречение.



5.3

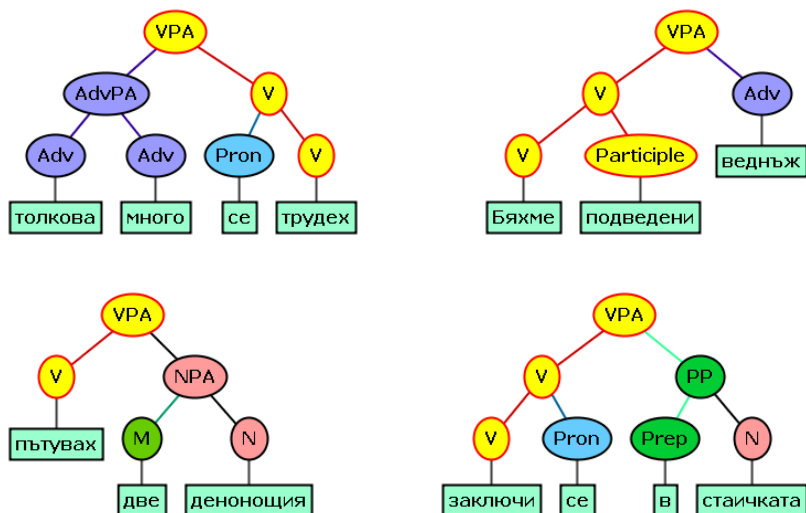
Глаголна фраза от типа опора-адюнкт – VPA

Фразите от този тип имат опора и един адюнкт. VPA може да има лексикална опора, ако тя не изисква комплемент и подлогът не е реализиран експлицитно. Следните лексикални и фразови елементи могат да бъдат опори: VPF, VPA, VPS, VPC, V, Participle, V-Elip, VD-Elip, CoordP и Verbalised. Т.е. всички възможни опори за VPS са възможни опори и за VPA, включително и самите фрази от типа VPS и VPA. Лексикалните елементи V, Participle и Verbalised са описани в раздела за лексикалните елементи по-горе. При наличие на няколко адюнкта към дадена лексикална опора, опората на една VPA фраза може да бъде друга VPA фраза. Трябва да обърнем внимание, че подобна рекурсия се появява само при този тип глаголни фрази³⁰. Фраза от тип VPF може да бъде опора на VPA фраза, ако запълващата фраза се реализира по-ниско от дадения адюнкт. Адюнктът може да бъде *съществително*, *фраза на съществителното*, *прилагателно* и *фрази на прилагателното*, *подчинено изречение*, *предложна фраза*, *адвербиална фраза*, *причастия*, *частици* или *координационна фраза*. Подобно на случая при глаголните фрази от тип опора-комплемент и опора-подлог, и тук опората може винаги да бъде определена еднозначно, тъй като нашето множество от категории е дефинирано така, че множеството на опорите и множеството на възможните адюнкти не се пресичат. Изключение от това правило е, когато опората и адюнктът са причастия, при което едното е в отношение на вторична предикация с неизразения подлог на другото. Да

³⁰ Когато има повече от един адюнкт, възниква въпросът за поредността на реализирането им. В момента ние сме приложили словореден критерий, според който: (1) ако адюнктите са еднакво близко до опората, те произволно се взимат един по един от нея и (2) ако адюнктите се различават по своята близост до опората, първо се взима най-близкият до нея. Разбира се, споменатото решение е с практически характер и не отменя изработването и прилагането на подходящ семантичен критерий според типа на адюнктите.

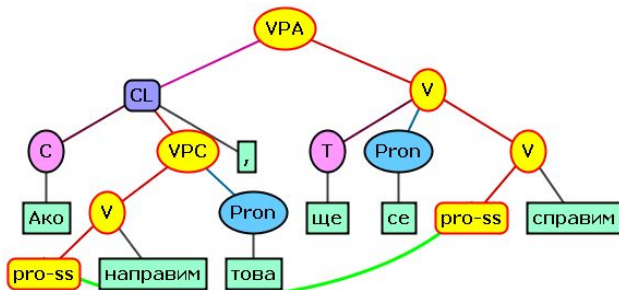
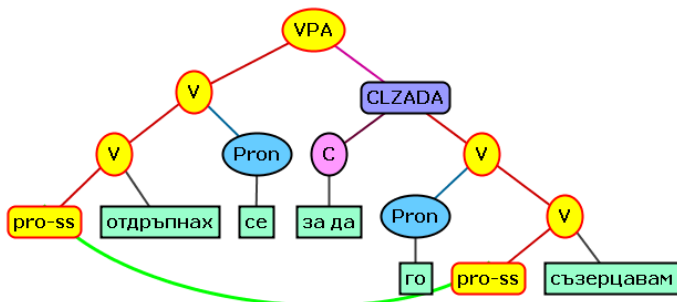
разгледаме например изречението *стоял онемял*. Смятаме, че този тип изречения имат най-често словоред, при който първа е лексикалната опора и след това идва адюнктът. Когато обаче опората е на втора позиция, тя е маркирана явно. Когато всички дъщери на една *VPS* фраза са координационни, то опората и адюнктът се определят на базата на категориите на аргументите на съответните координации.

Адюнктът се реализира в произволна словоредна подредба спрямо опората. Той може да се появи словоредно извън обсега на *VPA* фразата и тогава се маркира локално като *nid* елемент. Адюнктът може да се реализира и като *DiscA* елемент между опората и подлога или между опората и комплементите, или между подлога и комплементите. В някои случаи, когато адюнктът е наречие, адвербиална фраза или предложна фраза, той може да се реализира и в рамките на лексикалната опора. Най-често стои между спомагателния глагол *съм* и страдателното причастие. Например, в изречението *петното е напълно изтрито* наречието *напълно* е адюнкт. Адюнктът, участващ в отношение на вторична предикация, се съгласува с подлога или с някой от комплементите на лексикалната опора по число, род и лице. Опората и адюнктът могат да се реализират дистантно. Те могат да бъдат разделени от *DiscE* елементи, които са конституенти на подлога или на комплементите на същата лексикална опора, или *Pragmatic* елементи като парентетични изрази, звателни форми и други. Адюнктът се свързва с опората на базата на своя **MOD** атрибут. Той, от своя страна, се свързва по подходящ начин със синтактико-семантичните характеристики на опората. Следват примери на фрази от тип опора-адюнкт:

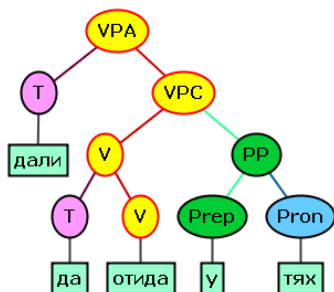


Първите два примера показват адвербиални адюнкти за количество и степен. Третият пример показва адюнкт, който е именна фраза за време. Последният пример е на адюнкт за място, изразен с предложна фраза.

Подчинените изречения, които могат да бъдат адюнкти, са от различен характер, например за цел и условие. В следващите графични представяния първият адюнкт е подчинено изречение за цел – *CLZADA*, въведено със съюза *за да*. Вторият адюнкт е подчинено изречение за условие, което е маркирано като подчинено изречение от общ вид – *CL*. В случаи като тези, когато има корелентна релация между неизразените подлози на подчиненото и главното изречение, представяме тази релация чрез връзка между *pro-ss* елементи.



Специален вид *VPA* са тези, които имат като адюнкт въпросителни частици: *нали*, *дали*, *нима* и модалната частица *да не би*³¹. Тези частици са адюнкти във *VPA* фраза, тъй като според Принципа на семантиката адюнктната дъщеря носи семантична информация към фразата. Те модифицират опората си, като формират въпрос или изразяват модалност.



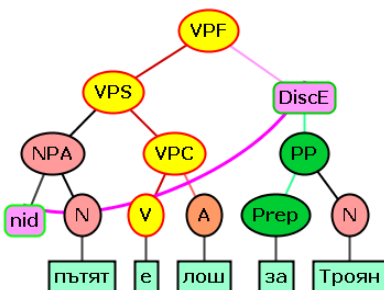
³¹ Някои от тях играят също така прагматична роля и тогава са маркирани с елемент *Pragmatic*.

5.4

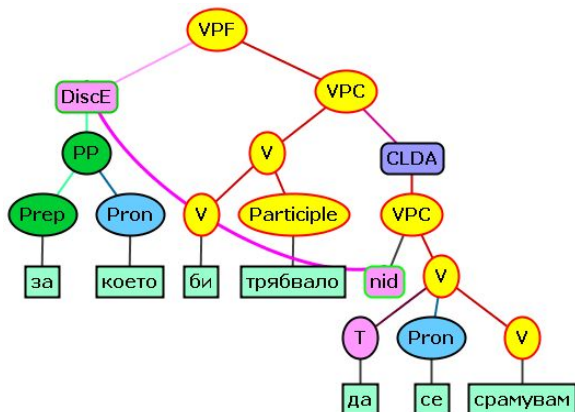
Глаголна фраза от типа опора-запълваща част – VPF

Фразите от този тип представят нелокалните (дистантните) реализации на конституентите. Те имат опора и една или повече запълващи части. *VPF* може да има само фразови опори: *VPA*, *VPS*, *VPC* и *CoordP*. Всички запълващи части са маркирани с функционалния елемент *DiscE*. Информацията за запълващата част се предава по йерархията на фразите от мястото, където съответната фраза липсва, до появата ѝ в рамките на *VPF* с помощта на атрибута **SLASH**, който съдържа информацията за липсващите конституенти.

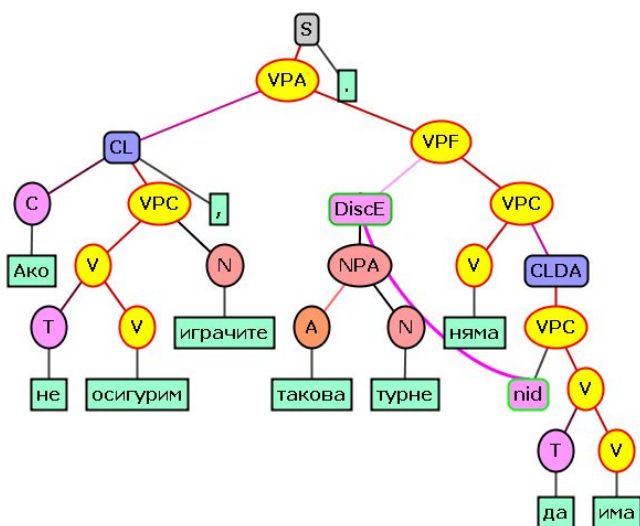
Запълващите части се реализират в произволен словоред по отношение на опората. Могат да бъдат дистантно разположени спрямо опората. Опората и запълващата част могат да бъдат разделени от прагматичен елемент, маркиран съответно като *Pragmatic*. Следва пример:



В примера има извлечена предложна фраза (*за Троян*), която е адюнкт в именната фраза-подлог на главното изречение. Изречението в обективен словоред е: *Пътят за Троян е лош*. Този анализ е реализиран на нивото на конституентната структура и ние го разглеждаме като анализ, който не използва следа според идеите на (Manning and Sag 1999). По този начин елементът *nid* указва на коя фраза е запълващата част, но не идентифицира словоредното място на запълващата част в рамките на фразата, от която е извлечена. Така в горния пример предложната фраза е задпоставена спрямо съществителното, но елементът *nid* е винаги първо дете на съответната фраза. Следващият пример е типичен за подчинените относителни изречения с безличен модален глагол, който има комплемент подчинено изречение. В тези случаи конституентът, съдържащ относителното местоимение, е разположен като първи конституент в словореда на подчиненото изречение. Много често относителният комплемент е извлечен от подчиненото изречение, където е зависима част (в примера долу рекцията на глагола е: *срамувам се за нещо*). В примера комплементът на 'да-изречението' е реализиран на нивото на модалния глагол. Всяка зависима част (комплемент, подлог, адюнкт) може да бъде извлечена от подчиненото изречение по този начин. Граматичната роля зависи от относителното местоимение.



Последният пример показва реализацията на запълваща част (DiscE) в рамките на фраза от тип VPA. Тук извлеченият комплемент от подчиненото изречение се реализира между VPC и адюнкта на по-високата фраза.



5.5

Подчинени изречения

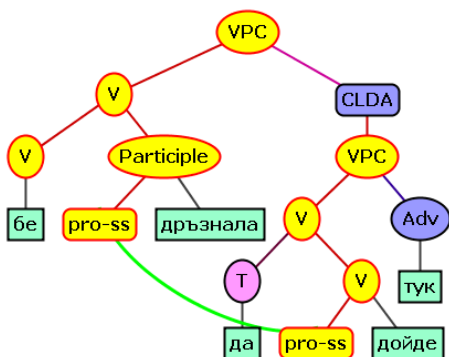
Глаголна фраза, която е реализирала всички зависими части на лексикалната си опора, е *наситена* (saturated) – и ще я наричаме *изречение*. Изреченията са главни или подчинени. Подчинените изречения могат да са комплети, подлози или адюнкти на други знакове. Подчинените изречения причисляваме към един (или повече) от следните типове: подчинени изречения с ‘да-комплекс’, маркирани с функционалния елемент CLDA; подчинени изречения,

въведени със съюза *че* – *CLCHE*; подчинени изречения, въведени със съюза *за да* или *да* със същото значение – *CLZADA*; въпросителни подчинени изречения – *CLQ*; относителни подчинени изречения – *CLR* и други подчинени изречения, маркирани с общ функционален елемент – *CL*. Типовете подчинени изречения са разграничени въз основа на типа въвеждаща дума (субординатор), а не на базата на функцията, която тази въвеждаща дума може да играе. Така например съюзът *че* може да въвежда следните изречения: подчинено допълнително, подчинено подложно, подчинени обстоятелствени за следствие или причина и др., но ние маркираме всички тях с функционалния етикет *CLCHE*. Функцията на подчиненото изречение се извлича от синтактичните етикети, които са под функционалния етикет – *VPC* посочва подчинено допълнително изречение, *VPS* посочва подчинено подложно изречение, а *VPA* посочва вид обстоятелствено подчинено изречение. За по-голяма яснота обаче тук ще обобщим поведението на подчинените изречения по тяхната граматична функция.

Подчинени изречения като комплементи към глаголна фраза

Това са изреченията тип *CLDA*, *CLCHE*, *CLQ* и по-рядко *CLR*. При типа *CLDA* някои глаголи от традиционните аналитични форми вече бяха посочени като опори, които изискват подчинено ‘да-изречение’ за комплемент – *цяха да дойде*. Тук ще споменем още една група глаголи – т.нар. *глаголи за контрол* (control verbs). При тях подлогът на опората е подлог и на глагола в подчиненото изречение. Примери за такива глаголи са: фазовите глаголи *започвам*, *продължавам*, *свършвам*, модалните *мога*³² и други – *готвя се*, *дръзнам*, *опитвам* и подобни.

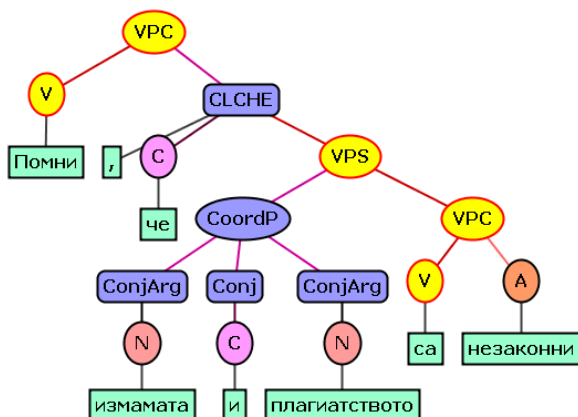
Комплементното ‘да-изречение’ е наситена фраза и следователно има изразен или неизразен подлог, който се корелира с подлога или някой от комплементите на главното изречение. Следва един такъв пример. Контролът на неизразения подлог на глагола в главното изречение върху неизразения подлог на глагола в подчиненото изречение е обозначен чрез корелиране на два *pro-ss* елемента. Срв. и с неграматичното **Бе дръзнала да дойдете тук*:



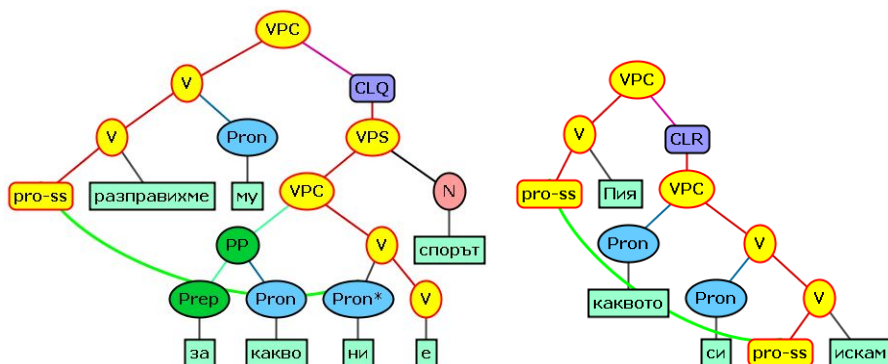
³² Модалните и фазовите глаголи в традицията на българските граматики образуват т.нар. съставно глаголно сказуемо. Аргументи в полза на идеята, че те образуват самостоятелно просто сказуемо привежда Пенчев в (Бояджиев и др. 1998, стр. 550-552).

Разбира се, има и случаи на ‘да-изречение’, при което опората на главното изречение е факултативно контролиращ глагол. Например *искам да отида*, но също *искам да отидеш*.

Изречения от тип *CLCHE* се избират често от глаголи за предаване на информация. Такива са: *твърдя, казвам, мисля, допускам, предполагам* и други. Следва пример:



За пълнота даваме примери с *CLQ* и *CLR*:



При примера с *CLR* може да се мисли за неизразен комплемент в главното изречение, към който *CLR* е адюнкт. Срв. *пия това, което си искам*. Но тъй като изпускането на комплемент в българския език не е граматикализирано явление, а по-скоро елипса, и при възстановяването на комплемента се налага изменение на релативната дума, предпочитаме да запазим традиционния анализ³³.

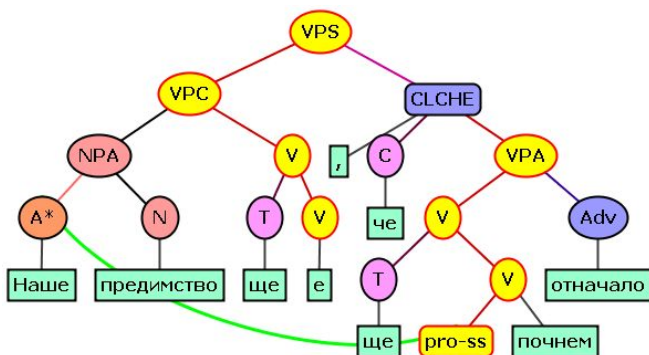
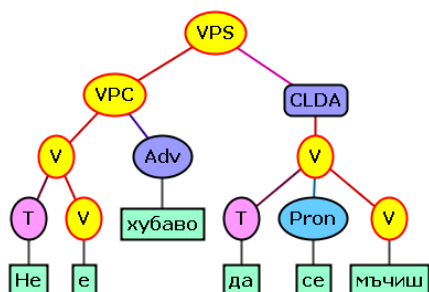
³³ Разбира се, една възможност е да въведем елементи, подобни на *pro-ss* за неизразения подлог, които да маркират неизразен комплемент (*pro-compl*) и неизразен адюнкт (*pro-adj*). Но за целта се изисква по-задълбочено проучване на условията за появата на последните два типа неизразяване.

Подчинени изречения като подлог към глаголна фраза

Според нас съществува йерархизиране на споменатите типове подчинени изречения при връзката им с главното изречение. За *CLDA* и *CLCHE* не се изисква кореферентност на аргументи на подчинените изречения с аргументи от главното изречение. За *CLQ* има главни изречения, които изискват кореференция между аргументи на главното и на подчиненото изречение, но има и такива главни изречения, за които това не е задължително. При *CLR* обаче винаги има кореференция между въвеждащата дума и част от главното изречение. В таблицата е посочено отношението на типа подчинено изречение към задължителността на кореференцията с главното изречение³⁴:

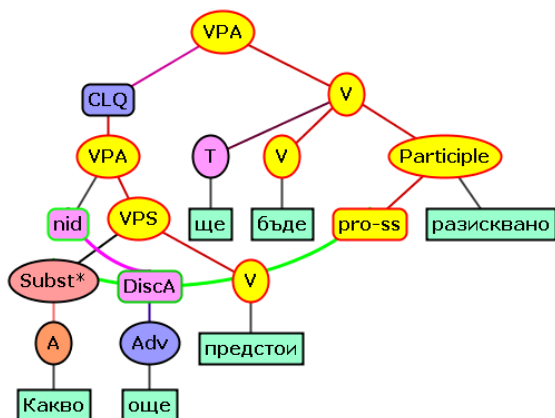
	CLDA и CLCHE	CLQ	CLR
Задължителна кореференция	-	+/-	+

Нека първо представим типовете *CLDA* и *CLCHE*. Двете примерни изречения долу могат да се представят чрез номинализации: *Това (= да се мъчиш) не е хубаво* и *Това (= че ще почнем отначало) ще е наше предимство*.



³⁴ Плюс означава задължителна кореференция, минус – незадължителна.

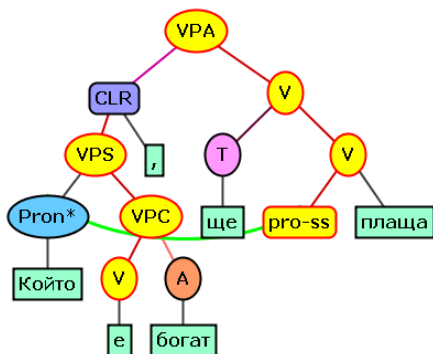
Типът *CLQ* в следващото изречение представя свързване на въпросителната дума (подлогът в *CLQ*) във въпросителното изречение с неизразения подлог на главното изречение.



Тъй като казахме, че *CLQ* може и да е свободно от корелираща релация, представяме един такъв пример: *попитаха къде отиваме*.

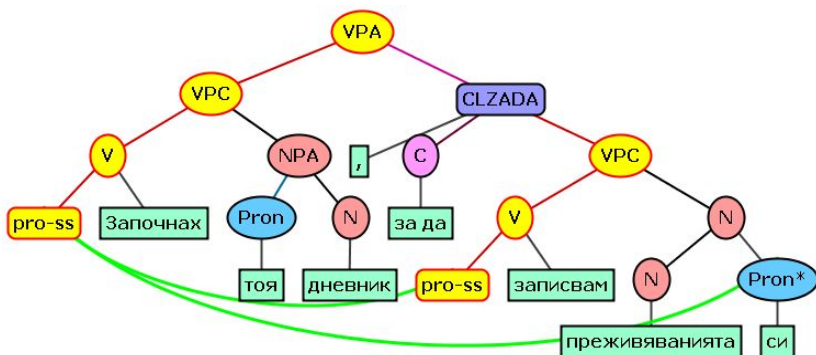
Типът *CLR* е интересен с това, че поставя под съмнение употребата си във функция на подлог. В граматиките е възприета идеята, че когато релативното изречение пояснява изразен подлог в главното изречение, то има функция на определително изречение (*Който пее, мой зло не мисли*). Но когато липсва изразен подлог в главното изречение, релативното се определя като подложно (*Който пее, зло не мисли*). Според нас при тези случаи е пренебрегната една основна синтактична особеност на българския език, а именно – че той е нулевосубектен език. При това положение функцията на релативното изречение и в двата случая е на адюнкт.

Разликата е, че в първия случай пояснява изразения подлог, а във втория – неизразения (повече по въпроса вж. (Osenova 2007)). Затова всички изречения от този тип ще бъдат представени само като адюнктни. По аналогичен начин са анализирани и тези въпросителни подчинени изречения, за които корелирането е задължително. Вж. примера по-горе за въпросителните изречения и следващия пример за относителните.

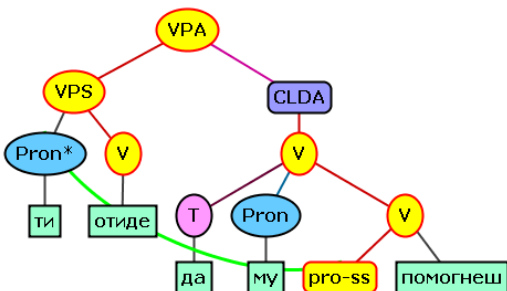


Подчинени изречения като адюнкт към глаголна фраза

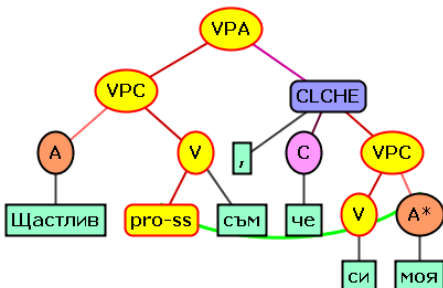
Изреченията, които могат да бъдат адюнкт към глаголна фраза са: *CLDA*, *CLZADA*, *CLCHE*, *CLQ* и *CLR*. Типът, който изпълнява само тази функция, е *CLZADA*. Подчиненото изречение е за цел. В примера е показана също така и кореферентно отношение между неизразените подлози на глаголите в главното и подчиненото изречение и възвратнопритежателната клитика:



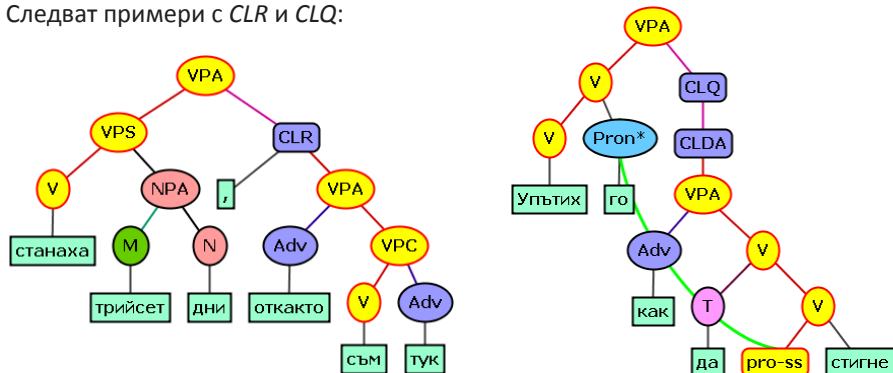
Синоним на *CLZADA* може да бъде и *CLDA*. Следва един пример, в който подлогът на главното изречение се корелира с неизразения подлог в подчиненото изречение:



В следващия пример *CLCHE* е в ролята на адюнкт за причина:

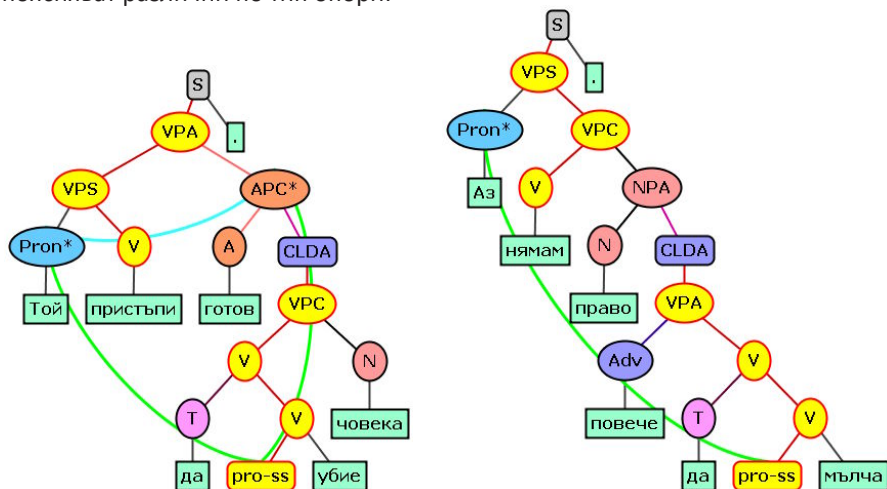


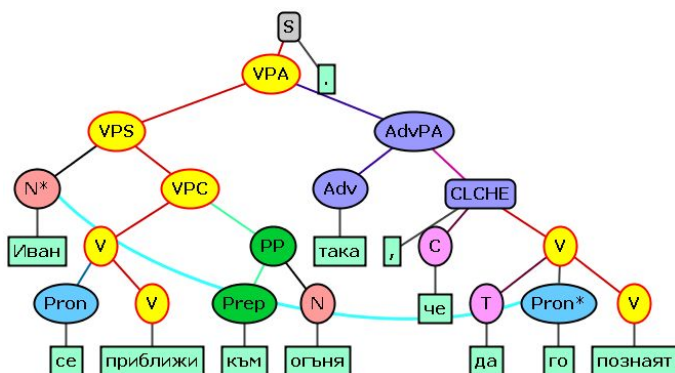
Следват примери с *CLR* и *CLQ*:



Подчинени изречения към други опори

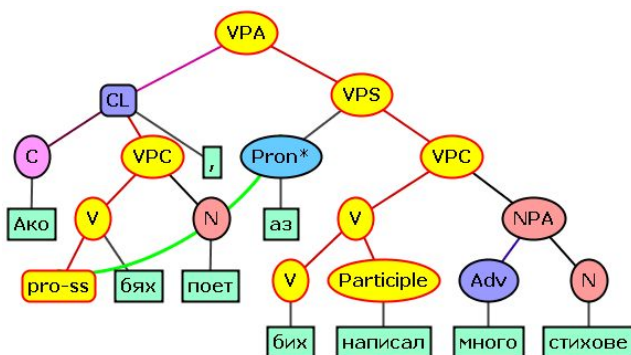
Освен като зависима част в глаголни фрази, подчинените изречения могат да поясняват и други типове фрази, като например адвербиални фрази, фрази на прилагателното, номинални фрази. Тук даваме пример с *CLDA* и *CLCHE*, които поясняват различни по тип опори:





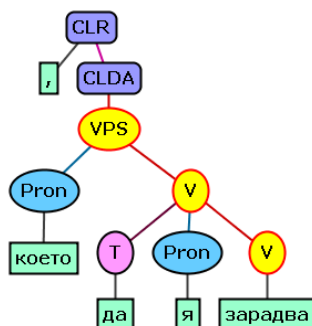
Други типове подчинени изречения – CL

Останалите видове подчинени изречения са представят чрез етикета CL. Тук даваме примери на най-честия случай, който е подчинено условно изречение. Това изречение се въвежда обикновено чрез подчинителния съюз *ако*.



Комбинирани подчинени изречения

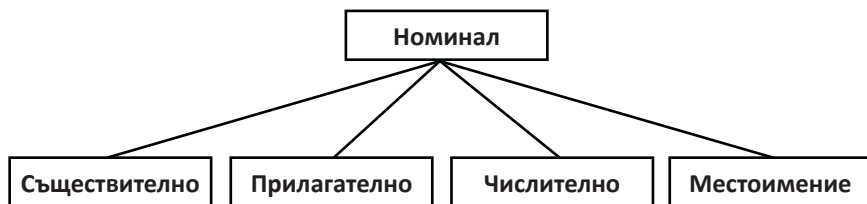
Случва се едно подчинено изречение да се отнася към повече от един тип. Това най-често става, когато 'да-изречение' е също така и въпросително или относително изречение. В тези случаи изречението първо се маркира като CLDA и след това като CLR или CLQ. Следва един пример, в който релативната дума има роля на подлог към 'да-изречението':



Именна фраза (номинална фраза)

От структурна гледна точка номиналната област съдържа лексикалния елемент съществително (*N*) и фразовите елементи от типа опора-комплемент *NPC*(omplement) и опора-адюнкт *NPA*(djunct). В тази част представяме фразовите елементи. Следвайки общата йерархия на реализация на зависимите части, получаваме следната последователност: *лексикален елемент съществително* → *именна фраза опора-комплемент* → *именна фраза опора-адюнкт*. Това означава, че първо се реализират елементите на лексикално ниво, после фразите от типа опора-комплемент и накрая фразите от типа опора-адюнкт. Разбира се, опората може да бъде изразена и с други именни лексикални елементи, като например местоимение, фамилно име, субстантивиран или номинализиран елемент.

От теоретична гледна точка приемаме, че в българския език съществителното без модификатори, квантификатори или детерминатори е функционално завършена именна фраза и се разглежда като фраза от типа 'само опора'³⁵. Уместно е да дефинираме йерархия от типове, в която типът *номинал* (nominal) включва подтиповете: съществително, прилагателно, числително и местоимение. Причините за това са следните: първо, всички именни елементи имат характеристика детерминация, което ги обединява, и второ, така се обхващат случаите на субстантивация:



Фразите от типа *опора-адюнкт* имат като зависима част не само модификатори, но също така детерминатори и квантификатори. В анализа си използваме механизма на **MOD** характеристиката, чрез която зависимата част избира своята опора. Нека припомним, че има и алтернатива – т.нар. „взаимна селекция“ между спецификатори и опори. Ние обаче разглеждаме и спецификаторите като модификатори (по-подробно вж. в (Осенова 2007)).

Фразите от типа *опора-комплемент* се отнасят за именни фрази от типа NN, в които първият елемент най-общо означава количество (*литър мляко*, *кошница ябълки*, *група студенти*) (по-подробно вж. в (Осенова 2004)). В този тип

³⁵ Това означава, че именната фраза се състои само от съществително опора и е без модификатори. Напр. *жена* или *столове* са фрази от този тип.

фрази действа Валентният принцип. В аргументната структура на опорното съществително има *NP*, което се изисква на семантично ниво. Но в списъка на комплементите, който е част от валентния списък и се реализира на синтактично ниво, този аргумент може да се появи или да не се появи. Срв. примерите: (1) *Групата пристигна* и (2) *Групата студенти пристигна*. Лексикалните спецификации на думата *група*, които показват отношението между аргументната структура и валентния списък, са следните:

<i>word</i> PHON	<група>
SYNSEM LOC CAT	[HEAD noun]
ARG-ST	[VAL COMP <>]
	< NP >

<i>word</i> PHON	<група>
SYNSEM LOC CAT	[HEAD noun]
ARG-ST	[VAL COMP <NP>]
	< NP >

И в двете лексикални спецификации аргументната структура съдържа изискване за *NP*, което да покаже обектите в групата. Но в първия вариант няма синтактична реализация на изискваното *NP* (изречение (1)), докато във втория има (изречение (2)). Затова списъкът на комплементите в първата схема е празен, а във втората съдържа *NP*.

Според представената идея опорните съществителни първо взимат комплементите си и после модификаторите си. Ето пример в скобов запис: [*NPA* *голяма* [*NPC* *група* [*NPA* *добри студенти*]]].

Макар че отглаголните и релационните съществителни са добри кандидати за образуване на тези фрази и че подобно решение е също подходящо за тях, в нашата граматика предпочитаме да дадем предимство на влиянието на опората, която е съществително. Така тези фрази са причислени към типа *опора-адюнкт*. Подобно на (Butt et al. 1999, p. 46) смятаме, че близостта на зависимите части с глаголната и именната опора се осъществява на нивото на тематичните роли, а не на граматичните функции. Напр. в изречението *Иван идва* съществителното *Иван* е подлог. В израза *идването на Иван*, *Иван* е част от адюнкт, но и в двата случая е с тематична роля на агенс.

Нека също си припомним, че във фразата *опора-адюнкт* адюнктните дъщери се взимат една по една, докато във фразата *опора-комплемент* комплементните дъщери се избират едновременно от опората.

Следва представяне на подтиповете на основните фрази: *опора-комплемент* и *опора-адюнкт*.

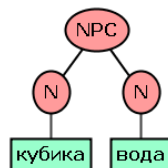
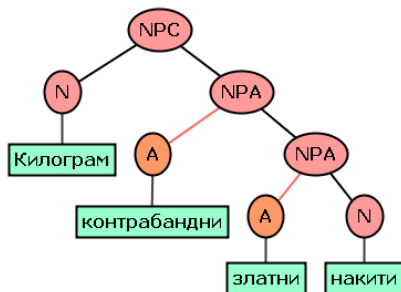
6.1

Именна фраза от типа опора-комплемент – NPC

Елементът с етикет *NPC* съответства на фразата от тип *опора-комплемент*. За такава фраза е задължително опорната дъщеря да има в аргументната си структура именна фраза за потенциален комплемент. Дали този аргумент ще се реализира, зависи от комуникативната ситуация (вж. по-подробно в (Ницолова 1984)). Опората на тези фрази може да бъде само лексикална и се заема от: *N*, *Subst*, *Nomin*, *N-Elip*, *ND-Elip*³⁶. Опората винаги е първият именен елемент, който обозначава количество. Следват описания на семантичните подтипове.

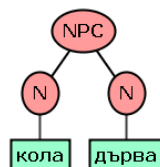
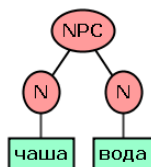
Количество – Субстанция

Опората на тези фрази изразява количество, а комплементът изразява същината, която се измерва. В първия пример опорното съществително (*килограм*) взема като комплемент цяла фраза от типа опора-адюнкт (*контрабандни златни накити*), а във втория пример опорното съществително (*кубик*) взема като комплемент лексикален елемент съществително (*вода*).



Съдържащо – Съдържимо

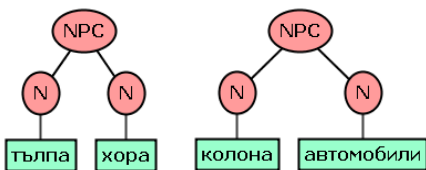
Това са фрази, в които опората означава нещо, което вмества и съдържа други същности, а комплементната дъщеря е съдържанието. В графиките са дадени класически примери за този вид фрази:



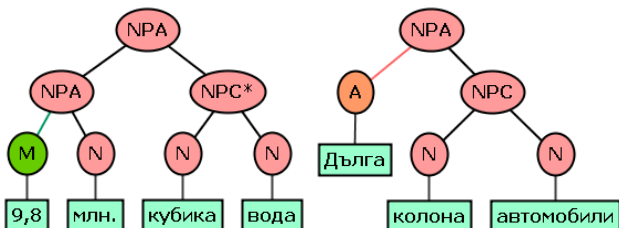
³⁶ Подобно на *V-Elip* и *VD-Elip*, *N-Elip* означава елипса на съществително, което се възстановява в рамките на изречението. *ND-Elip* маркира елипса на съществително, което се възстановява в рамките на по-широкия контекст.

Форма на групиране – Същини

В тези фрази опорното съществително означава събирателност на обекти, а комплементът означава обектите, които 'се събират' в едно множество. Напр. *група студенти, вид сирене*³⁷.



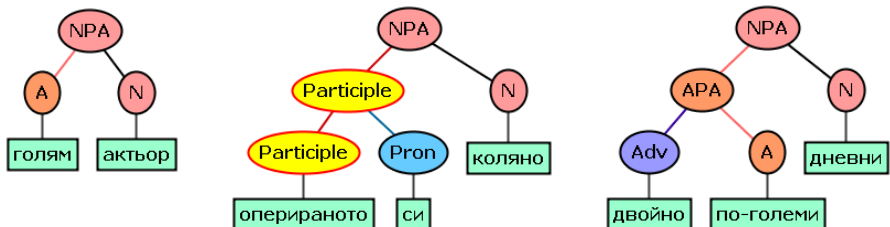
Когато фразите *опора-комплемент* се модифицират от адюнкт(и), адюнктите се добавят последни в именните проекции. Т.е. опората взема първо комплементите си и след това се модифицира от адюнкти.



6.2

Именна фраза от типа опора-адюнкт – NPA

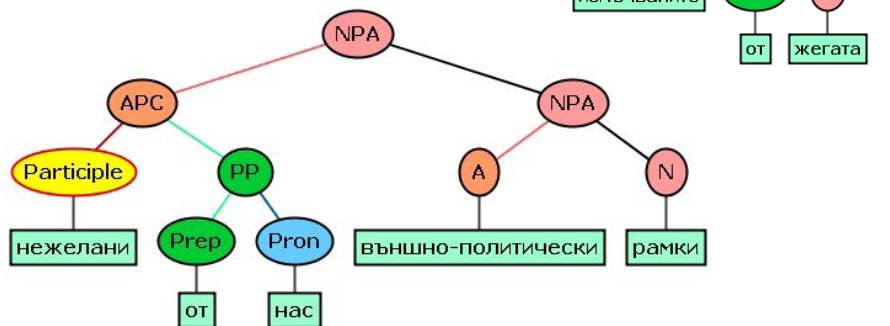
Както беше посочено по-горе, елементът *NPA* съответства на именна фраза от типа *опора-адюнкт*. Всяка фраза от този тип трябва да има опорна дъщеря и адюнктна дъщеря. Опората може да бъде лексикален (*N, Pron, H, Subst, Nomin, N-Elip, ND-Elip, CoordP*) или фразов елемент (*NPA, NPC, N-Elip, ND-Elip, CoordP*). Адюнктът съответно може също да бъде лексикален (*N, M, A, Pron, Participle, CoordP*) или фразов елемент (*AP, PP, NPA, CLDA, CLQ, CLR, CoordP*). Опората в този тип фрази се определя безпроблемно в случаите, когато адюнктната дъщеря не е съществително. В случаите на фрази от типа *NN* обаче опората може да бъде както първото (*гора закрилница*), така и второто съществително (*злато жена, дядо Петър*). Когато опората е вторият конституент, този факт се маркира експлицитно със знака *+/* при етикета на възела на фразата. Следват примери на типовете фрази с различни видове опори и адюнкти:



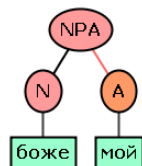
³⁷ Тук нямаме възможност да дискутираме по-подробно разликите между тези подтипове. Повече информация за семантичните отношения в именните фрази от типа *NN* може да се намери в (Алмалех 1993).

Първият пример показва основния случай на съществително, модифицирано от прилагателно. Вторият пример демонстрира адюнкт, който е лексикален елемент и включва както причастие, така и възвратнопритежателна клитика към него. Третият пример показва съществително, което е модифицирано от фраза на прилагателното (APA).

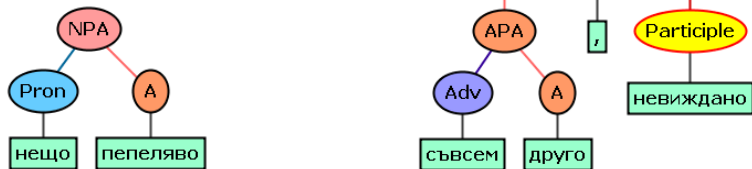
Следват примери, в които лексикалният елемент съществително име и фразовият елемент NPA са модифицирани от фраза тип *опора-комплемент*, в която опората е причастие.



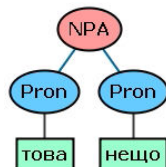
Опорно съществително в звателна форма най-често е пре- или постмодифицирано от прилагателно или местоимение. Примерът показва постмодификация на съществителното с притежателно местоимение:



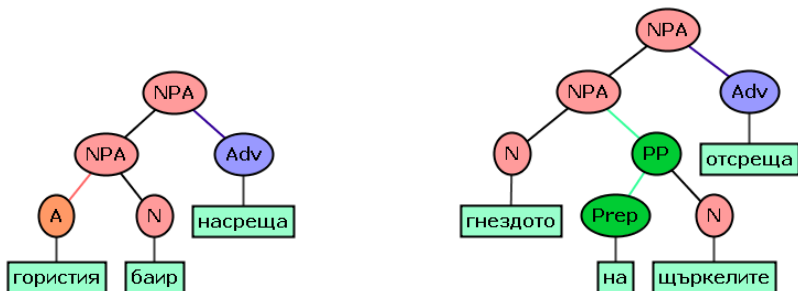
Неопределителното, въпросителното, отрицателното субстантивно местоимение може да бъде постмодифицирано от прилагателно или фраза на прилагателното. В първия пример по-долу неопределителното местоимение *нещо* е постмодифицирано от прилагателното *пепеляво*. Във втория пример опора отново е *нещо*, но вече постмодифицирано от координационна фраза, чиито конюнкти са фраза на прилагателното (*съвсем друго*) и причастие (*невиждано*). Координационната връзка в случая е запетаята.



Неопределителното и отрицателното субстантивно местоимение могат да бъдат премодифицирани от подходящи адюнкти (повече за поведението на *нещо* и *нищо* (вж. Пенчев 1962)). В този пример показателното местоимение *това* премодифицира *нещо*.



Друг случай е, когато съществително или именна фраза са постмодифицирани от наречие. И в двата примера, които следват, наречието постмодифицира именни фрази с различна степен на сложност:



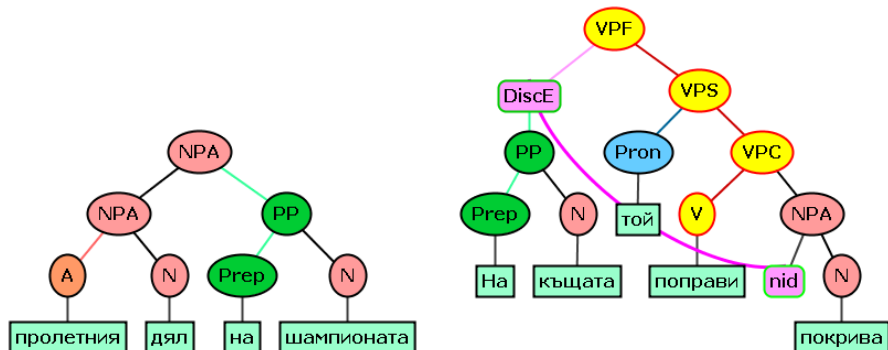
6.3

Типове адюнкти в именните фрази

Предложни фрази

Именната фраза от типа 'само опора', както и другите два типа именни фрази (опора-комплемент и опора-адюнкт) могат да се постмодифицират от предложна фраза – *PP*. Приемаме, че предложните фрази се реализират след типичните премодифициращи фрази³⁸ (прилагателни, фраза на прилагателното и др.). В първия пример по-долу е даден типичен случай на последователна реализация на пре- и постмодифицираща фраза. Вторият пример обаче предлага по-интересно езиково явление. То се нарича 'извличане на предложна фраза от именна фраза'. Даденият по-долу втори пример в обективен словоред е следният: *Той поправи покрива на къщата*. Тук предложната фраза модифицира опорното съществително *покрива*. Но в графичното представяне е дадено възможното разместване на зависимите части в изречението. Предложната фраза е маркирана като елемент, който е 'откъснат' словоредно от опората си (*DiscE*), но конституентната му позиция на адюнкт в именната фраза е посочена чрез специалния елемент *nid*.

³⁸ Разбира се, това решение е условно и гарантира непротиворечивост и кохерентност на анализите. Но еднакво вярно е предложната фраза да се реализира преди премодификаторите. Ситуацията се дължи на наличието на т.нар. изкуствена многозначност (*spurious ambiguity*). С други думи, от лингвистична гледна точка няма значение коя зависима част ще се вземе първа от опората. Разсъждения по този въпрос има в (Пенчев 1993).



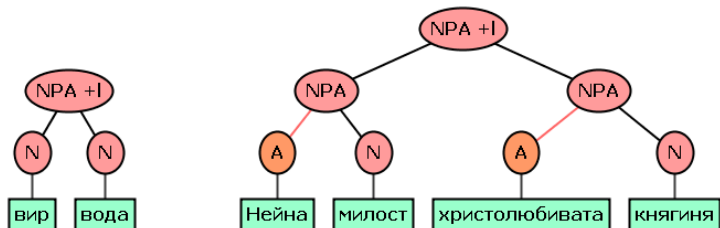
Самостоятелни съществителни имена и именни фрази

В апозиция (фрази от семантичния тип *същина-квалификация*). От словоредна гледна точка съществуват два основни типа фрази – адюнкт в препозиция и адюнкт в постпозиция. До този момент беше лесно да се изведе кой елемент е опората във фразата само на базата на морфологичния, лексикалният етикет или словоредът. Но в рамките на апозициите това невинаги е възможно. Затова се налага да маркираме експлицитно коя е опората в определени случаи. Въвеждаме тип опора с три характеристики: (а) опора по подразбиране – *default* (default value). В този случай се използват правилата за определяне на опората (напр. *Ако в една фраза има лексикален етикет N, последван от предложна фраза PP, то опората на именната фраза е N*); (б) опора в първа словоредна позиция – *f* (first). В този случай опората е първият елемент на именната фраза. Този случай е приет по премълчаване; (в) опора в последна словоредна позиция – *l* (last). В този случай опората е последният елемент във фразата.

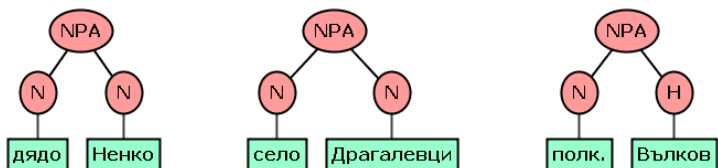
Има два типа фрази, в които модифициращото име е първо: (а) фрази от типа на *злато жена* и (б) фрази от типа на *дядо Петър, професор Тодоров, град Плевен*. Първият тип е от т.нар. *изафетни конструкции*³⁹, в които едно съществително модифицира друго. Този тип се активизира напоследък – срв. *огън кафе, бизнес среща* и подобни. Вторият тип представя следния модел: семантичен клас адюнкти – звание, професия, роднински названия и подобни, последвани от опора-собствено име. Разбира се, в този модел опората се разпознава лесно по информацията за собствено име в морфологичния етикет⁴⁰.

³⁹ По-подробно в (Байрамова 1982).

⁴⁰ В родната и чуждата езиковедска литература има спорове кое точно съществително е опора в апозитивни конструкции. **Целта ни тук е да представим модела на отношението опора-адюнкт без оглед на конкретното решение.** В нашата интерпретация приемаме собственото име за опора.



И двата примера горе показват проблема с идентификация на опората. Затова в дясната част на максималната проекция на фразата е указан признакът +I, т.е. опората е вторият елемент.

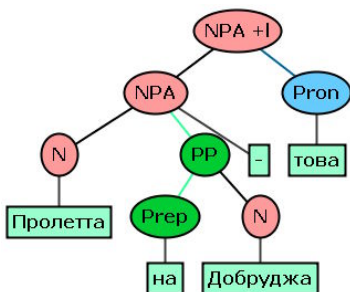


Горните примери не са проблемни по отношение на опората, тъй като в морфологичния етикет има информация, че вторият елемент е съществително собствено име. Следователно той е опората на фразата.

Следват примери, в които адюнктната дъщеря е на втора позиция и опората е първият елемент. По премълчаване първото съществително е опората.



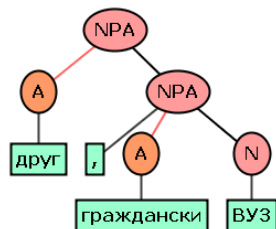
Фрази с опора местоимение и адюнкт – именна фраза. В случаите, когато местоимение номинализира цяла фраза, приемаме местоимението за опора. В примера, който следва, показателното местоимение *това* 'замества' фразата *пролетта на Добруджа* и проектира максимална именна фраза от типа *опора-адюнкт*:



Препозитивна модификация от особен тип

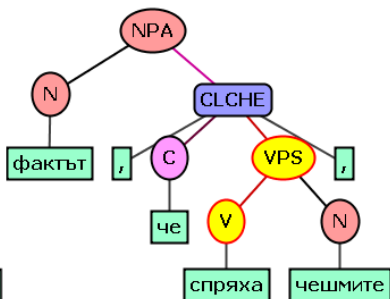
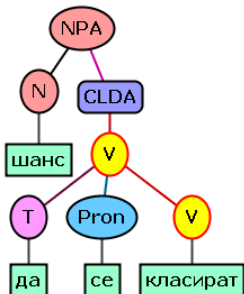
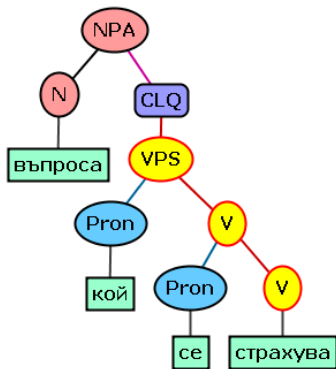
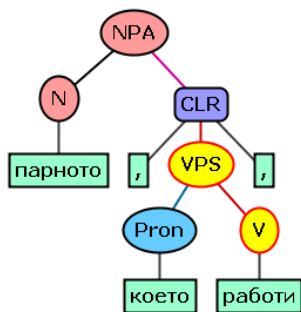
Има случаи, макар и рядко, когато препозитивните адюнкти са разделени от запетая, но тя не изразява координационни отношения между адюнктите.

В примера, който следва, фразата не е синонимна нито на координационна структура – срв. *друг и граждански ВУЗ*, нито на класическа субординационна структура – срв. **друг граждански ВУЗ*. Тя има субординационно доуточняващо значение – срв. *друг, а именно (с други думи) – граждански ВУЗ = друг ВУЗ, който е граждански*.



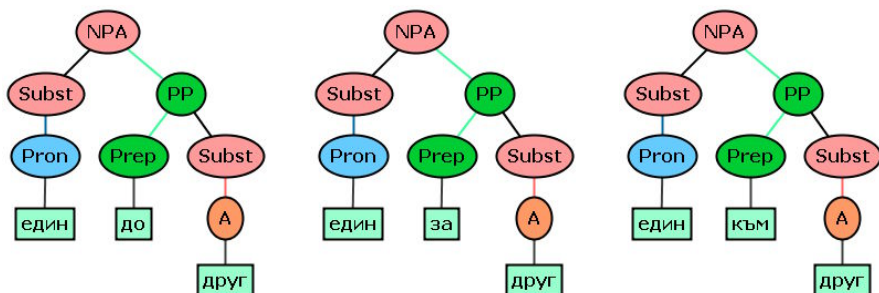
Подчинени изречения

Типични постмодификатори, подчинените изречения се въвеждат с: относително местоимение (CLR), въпросително местоимение (CLQ), да-комплекс (CLDA) и съюза *че* (CLCHE). Те могат да модифицират както опората, така и нейните други модификатори. Дадени са графични примери с всеки тип подчинено изречение от изброените. Първият пример представя именна опора, модифицирана от подчинено относително изречение. В тези случаи относителното местоимение винаги е кореферентно в пълна или частична степен с именната си опора от главното изречение. Вторият пример представя именна опора, модифицирана от въпросително местоимение. За разлика от предходния случай, тук липсва кореференция между името и въпросителната дума. Аналогично е за изреченията, въведени с *да* и *че*. Във всички представени типове именната опора номинализира част от съдържанието или цялото съдържание на подчиненото изречение. Затова, с изключение на изреченията с относително местоимение, има семантични ограничения върху типа на опората. Напр. тя трябва да обозначава факти от когнитивно или емоционално естество: *въпрос, мисъл, мнение, факт, желание, страх* и др.



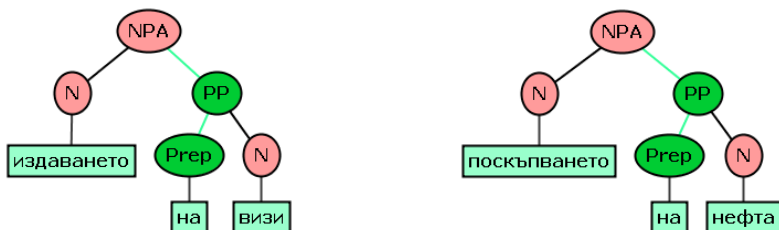
Реципрочни изрази от типа 'един-друг'

Нашата интерпретация на реципрочните изрази е следната. Първият елемент *един* се разглежда като субстантивиран и след това се модифицира от другия елемент, който може да бъде: самостоятелен (*Те се виждаха **един друг***) или комплемент в предложна фраза (*Те си помагаша **един на друг***). Субстантивацията на втория реципрочен елемент не се изисква от предлога, но заради единността на анализите е добре той също да бъде представен като субстантивиран⁴¹.



Отглаголни съществителни

Отглаголните съществителни наследяват аргументната структура на производния си глагол, но с модификации в конституентното изразяване (често комплементите се изразяват с предложни фрази) и задължителността им. Затова, както беше отбелязано в частта за лексикалните елементи, ние ги анализираме като адюнкти към именната опора:



Ако съпоставим горните примери със съответстващите им глаголни опори, се получава следното: *визи* е комплемент на глагола *издавам* (*издавам визи*), а *нефт* е подлог на глагола *поскърпвам* (*нефтът поскърпва*). Но и двата аргумента са изразени чрез предложни фрази към именните си опори.

⁴¹ Причината, поради която думата *друг* не се разглежда като местоимение в тази конструкция е, че като речникова единица тя е в реда на прилагателните. Освен това при процеса на субстантивация двата лексикални елемента стават равностойни. Затова не смятаме за нужно да променяме нейния лексикален етикет в подобни изолирани употреби.

7

Фраза на прилагателното (адективна фраза)

В граматиката разграничаваме два типа фразови елементи с опора прилагателно: *опора-комплемент* – APC(*omplement*) и *опора-адюнкт* – APA(*djunct*). Спазвайки реда на общата реализация на зависимите части, първо се реализира лексикалният елемент, после фразата *опора-комплемент* и накрая фразата *опора-адюнкт*. Опората на фразата може да бъде изразена не само от прилагателно, но и от адективно местоимение или причастие.

7.1

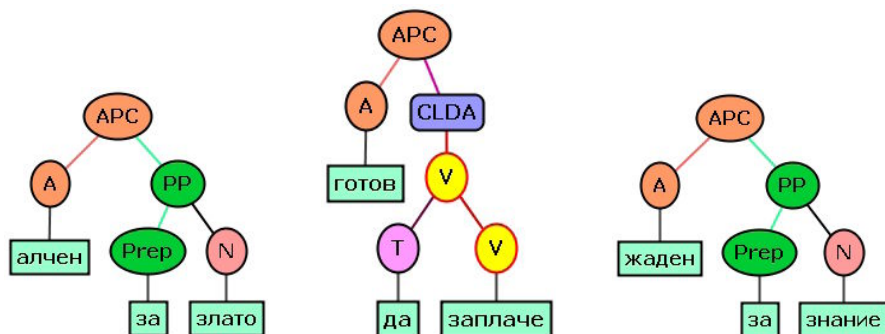
Фраза на прилагателното от типа опора-комплемент – APC

Тази фраза има за опора само лексикален елемент със следните етикети – A, Participle, CoordP.

Трябва да се има предвид, че прилагателното и причастието се различават в поведението си на опори. Причастието наследява аргументната структура на изходния глагол и затова появата и броят на комплементите зависи изключително от нея. От друга страна, прилагателното няма склонност да взима комплементи, с изключение на малък брой релационни прилагателни, чиято типология не е проучена изцяло. Нека разгледаме подтиповете на фразата APC:

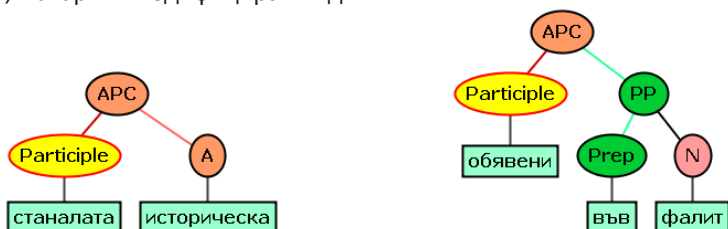
Опора прилагателно с комплемент(и)

Ето някои прилагателни, за които се посочва, че изискват комплементи: *алчен за*, *готов на/да*, *жаден за*. Обикновено комплементите са изразени чрез предложни фрази или подчинени изречения:



Опора причастие с комплемементи

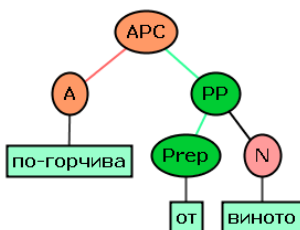
Както беше посочено вече, причастията наследяват аргументната структура на глагола, макар и в модифициран вид:



В първия пример глаголното съответствие би било *нещо става историческо*. Глаголът *ставам* тук е десемантизиран и е синонимен на спомагателния глагол *съм*. Припомняме, че в нашите анализи *съм* взима комплемементи. Вторият пример показва фраза, в която единият от двата задължителни комплемента липсва – *някой обявява нещо във фалит*. Липсва това, което е обявено във фалит.

Сравнителни и превъзходни форми на прилагателното

Сравнителните и превъзходните форми често изискват специфициране на същините, които те сравняват с оглед на определени характеристики. Приемаме същините, спрямо които се сравнява, за комплемементи. Обикновено комплементите се въвеждат с предлозите *от*, *сред* и др.

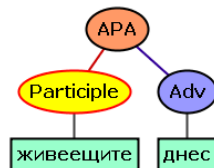
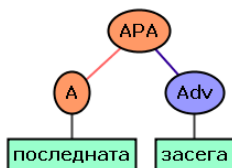
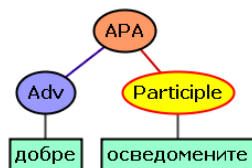


7.2

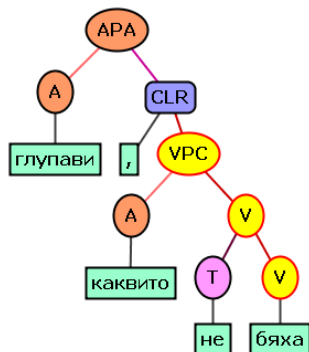
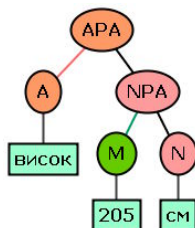
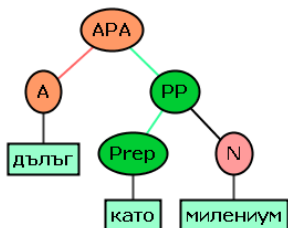
Фраза на прилагателното от типа опора-адюнкт – АРА

Тази фраза може да има за опора както лексикален елемент (A, Participle, CoordP), така и фразов (APA, APC, CoordP).

Адюнктите се изразяват чрез наречия, именни елементи, предложни фрази, подчинени изречения, координационни фрази. Те могат да бъдат препозитивни или постпозитивни. Ако опората е причастие, адюнктите на изходния глагол могат да се появят като адюнкти и в адективната фраза. Ето примери, в които адюнктът е изразен с наречие в пре- или постпозиция:

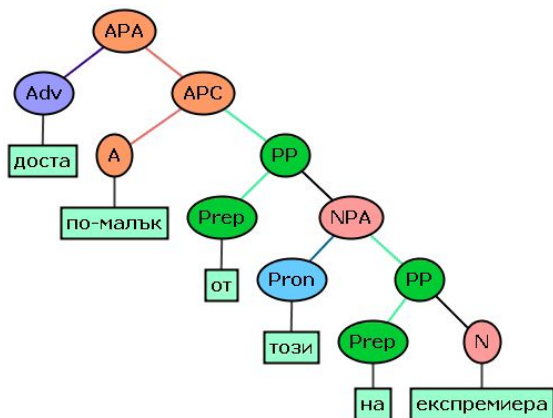


Адюнктите могат да се изразят и с подчинено изречение, именни елементи, предложни фрази:



В горните илюстрации са представени съответно: в първия пример – адюнкт предложна фраза, във втория пример – адюнкт именна фраза и в третия пример – адюнкт подчинено изречение с относително местоимение.

Срещат се, разбира се, случаи, в които опората на фразата *APA* е *APC*. В примера, който следва, сравнителната форма на прилагателното *по-малък* първо си взема комплементната дъщеря (*PP*) и после прибавя препозитивното наречие *доста*, като прави адюнктна фраза:



8

Фраза на наречието (адвербиална фраза)

Подобно на типовете фрази при прилагателното, при наречието като опора също има два типа фрази: опора-комплемент (*AdvPC(omplement)*) и опора-адюнкт (*AdvPA(djunct)*). Редът на реализация на фразите е същият като при разгледаните вече фрази. Лексикалната опора може да бъде освен наречие, също и местоименно наречие или деепричастие.

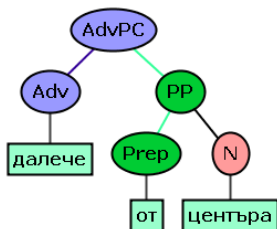
8.1

Фраза на наречието от типа опора-комплемент – AdvPC

Този тип фрази имат за опора само лексикални елементи: *Adv*, *Gerund*, *CoordP*. Подобно на причастията, деепричастието е глаголна форма и затова наследява аргументната структура на глагола. Следват подтиповете фрази според вида на опората.

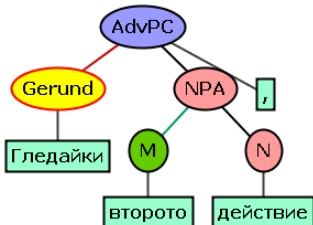
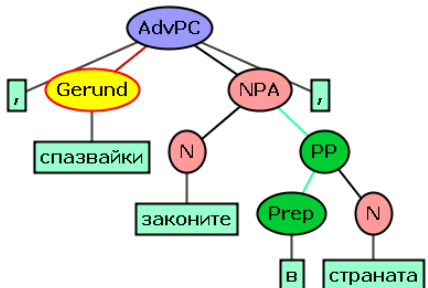
Опора наречие с комплементи

Някои наречия изискват комплементи. Например:



Опора деепричастие с комплементи

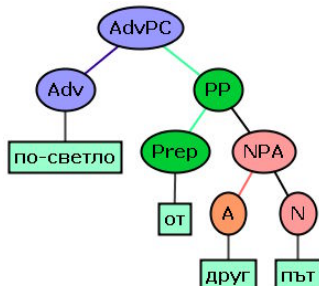
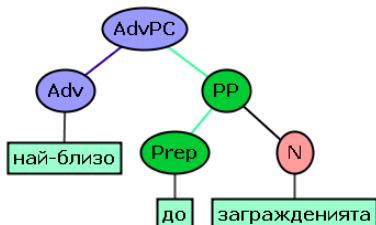
В примерите по-долу деепричастията изискват комплементите, които са в аргументната структура на изходната лична форма на глагола:



Фразите с деепричастие винаги са отделени пунктуационно от останалата част на изречението.

Сравнителни и превъзходни форми на наречието

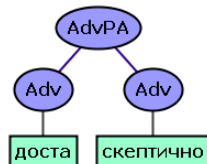
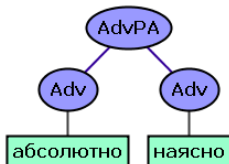
Подобно на прилагателните, степенуваните наречия изискват компленти:



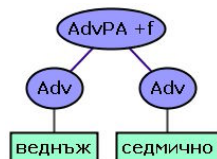
8.2

Фраза на наречието от типа опора-адюнкт – AdvPA

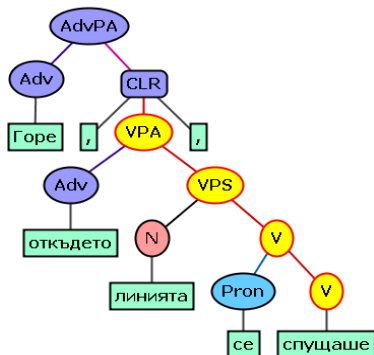
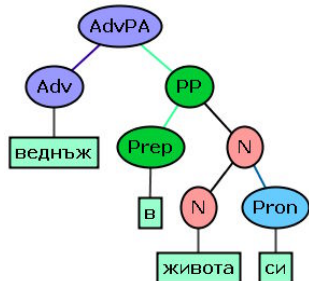
Този тип фрази могат да имат за опора както лексикален (*Adv*, *Gerund*, *CoordP*), така и фразов елемент (*AdvPA*, *AdvPC*, *CoordP*). Адюнктите могат да бъдат: наречия, именни елементи, предложни фрази, подчинени изречения, координационни фрази. Във фразите от типа *AdvAdv* обикновено първото наречие е адюнкт, а второто – опора:



Но има и случаи, при които опората е преди адюнкта. Тя е маркирана с характеристиката *+f*.



Следват примери, при които адюнктната дъщеря е изразена с предложна фраза или с подчинено релативно изречение:



9

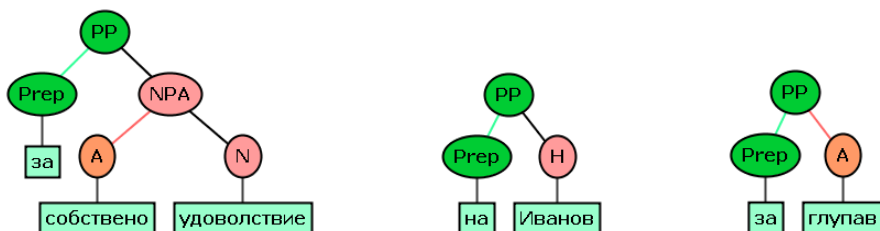
Предложна фраза

Предложните фрази биват два типа: *опора-комплемент* (PPC(omplement)) и *опора-адюнкт* (PPA(djunct)). Тъй като вторият тип се среща много рядко, не маркираме тези разлики експлицитно. Т.е. всички фрази се маркират с по-общия лексикален етикет *PP*, а особеностите се виждат от вътрешната структура на фразата.

9.1

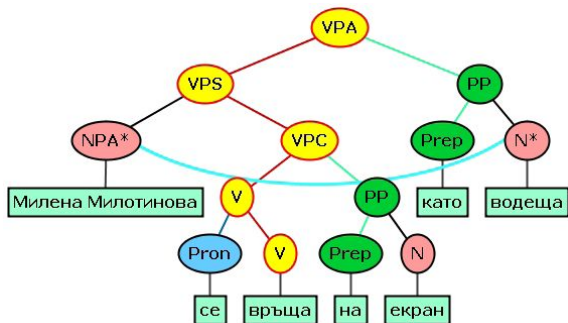
Предложна фраза от типа
опора-комплемент – PP(C)*Предложни фрази с именни комплементи*

Именните комплементи са най-честият тип предложни аргументи. Опората в тези фрази е предлогът. Той се маркира с лексикалния етикет *Prep*. Тъй като в по-широкия дискурс опората може да се изпуска, се въвежда етикетът *Prep-Elip*. Например, 'Радиацията в Косово била по-слаба от София.' се интерпретира като 'Радиацията в Косово била по-слаба от [N-Elip радиацията] [Prep-Elip в] София.' Комплементната дъщеря може да се изрази с лексикален или фразов елемент: *N*, *H*, *A*, *APA*, *Pron*, *Nomin*, *Subst*, *NPA*, *NPC*, *CoordP*. Комплементът е винаги словоредно след опората:



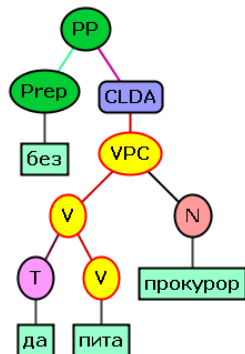
В горните примери са дадени различни елементи като комплементи на предлози. Това са съответно именна фраза (*NPA*), фамилно име (*H*) и прилагателно (*A*).

Има предложни фрази с именни комплементи, при които опорните предлози са *като* или *за*. В случаите, когато предложните фрази означават вторична предикация (т.нар. сказуемни определения), именните комплементи се корелират с подлога или допълнението. Ето пример за вторична предикация. Показана е корелацията между собственото име *Милена Милотинова*, което е подлог в изречението, и предложната фраза *като водеща*:



Предложни фрази с компленти подчинени изречения

Има предлози, които взимат подчинени изречения като компленти ('да-изречения', 'че-изречения' или и двете). Тези предлози са: *освен*⁴², *вместо*, *без*, *преди*. Бихме искали пак да обърнем внимание, че част от компонентите на традиционно разглежданите сложни съюзи (напр. *преди да*, *без да*) в нашия анализ се смятат за предлози, които взимат не само именни фрази, но и подчинени изречения като компленти. Т.е. ние разширяваме групата на езиковите елементи, които са аргументи на тези предлози. Това става само в случаите, когато е положителен тестът за вмъкване на лексикален елемент между частите на кандидата за сложен съюз⁴³. В този пример предлогът *без* взима 'да-конструкция' като комплемент.



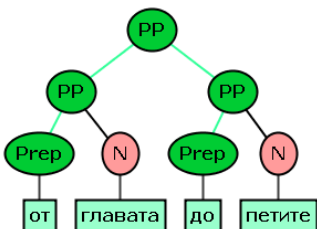
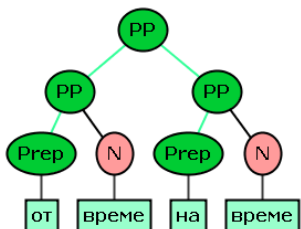
Предложни фрази с компленти предложни фрази

Тук се включват изрази за обхват или диапазон. Обикновено става дума за темпорален или пространствен интервал. Следват типове:

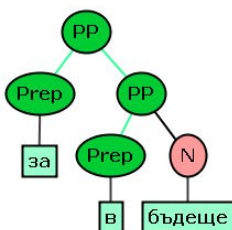
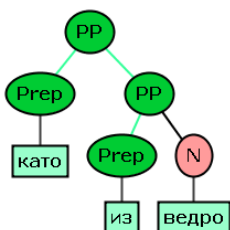
Идиоматични изрази от типа: *от време на време*, *от главата до петите*. В тези конструкции условно смятаме първата предложна фраза за опора, а втората за комплемент, тъй като без нея първата няма смисъл. Друга възможна алтернатива е тези конструкции да се разглеждат като конструкции без опора (*non-headed*).

⁴² Случаят с *освен* е интересен. Със сигурност е предлог в съчетанието *освен да*. Но за другите варианти се държи по-скоро като част от сложен съюз: *освен че*, *освен ако*.

⁴³ Разбира се, нашите решения могат да се приложат и по-радикално. Например, чрез следното правило: всички предлози, които са част от сложни съюзи и независимо от пропускливостта на лексикални елементи, да взимат подчинени изречения като компленти. Тогава и последователности като: *за да* и *след като*, които в момента са в списъка на сложните съюзи, могат да минат към модела *предлог+подчинено изречение*.

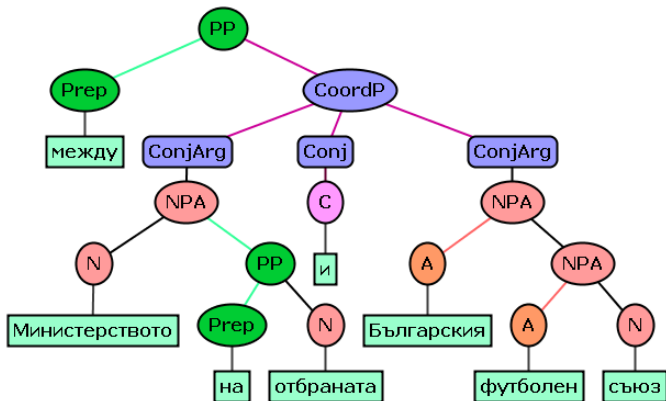


Предложни фрази с комплемент една предложна фраза: Напр. *като за теб*. Разбира се, има ограничения в съчетаемостта на предлозите. Следват примери:



Предложни фрази със събирателен комплемент

Става дума за фрази с опора предлозите *между* и *измежду*. Те изискват комплемент, който назовава повече от един обект. Този обект може да се изрази само с един лексикален знак: *Той е най-умен измежду тях*. Но също така често се изразява с координационна фраза:

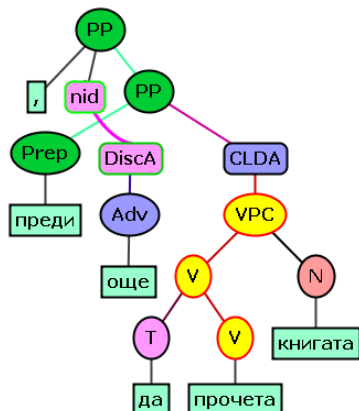
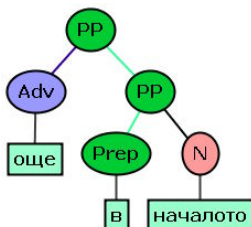


9.2

Предложна фраза от типа опора-адюнкт – PP(A)

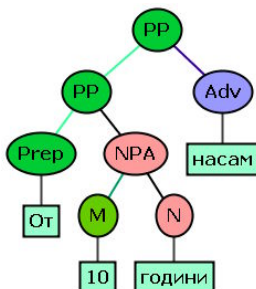
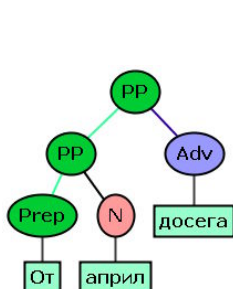
Предложни фрази с адюнкти – наречия

Това са предложни фрази с емфатично наречие. Типични емфатични наречия са: *още*, *само* и др. Опората винаги е предложната фраза:



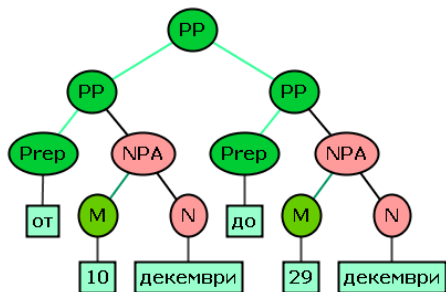
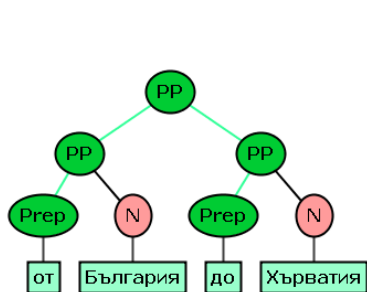
В първия пример има присъединяване на емфатичното наречие *още* към предложна фраза и проекцията отново е предложна фраза. Вторият пример представя по-сложно езиково явление, при което същото емфатично наречие се реализира словоредно между предлога *преди* и 'да-конструкцията'. Затова се активира механизмът за маркиране на дистантно разположени конstituенти, който показва, че конstituентната позиция на наречието е по-високо в дървото.

Два примера на предложни фрази с адюнкт – наречие:



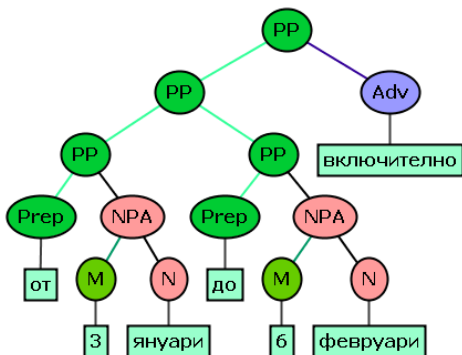
Предложни фрази с адюнкти – предложни фрази

Темпорални и пространствени изрази за интервал: *от Варна до Бургас*, *от 6 юни до 20 август* и др. Когато двете предложни фрази са една до друга, анализът е, както следва:



Когато обаче двете предложни фрази са разделени от друг лексикален материал, те се реализират една по една, тъй като това е начинът на реализация на адюнктите. Срв. *От България се стига бързо до Хърватия.*

Друг случай е, когато темпорален израз се модифицира от наречие. Тогава наречието се свързва най-високо в дървото:



Координационните фрази (етикет *CoordP*) съответстват на т.нар. *неопорни фрази* (non-headed phrases). В момента това е единственият подтип неопорна фраза в нашата граматика⁴⁴. Следствието от подобен анализ е, че координацията представлява ‘плоска структура’ (flat structure) на координиращи елементи⁴⁵ (етикет *Conj*) и координационни аргументи (етикет *ConjArg*). Нека напомним, че всички елементи са разположени на едно ниво. От гледна точка на участието на координацията като конституент в структурата на други фрази, тя трябва да дефинира знак, който, от една страна, да бъде избиран или да избира, а от друга страна, да съвместява различните особености на своите аргументи и координиращи елементи. Нашият анализ е основан на идеята, че аргументите на координацията трябва да съвместяват своите граматични функции, докато синтактичните им категории могат да бъдат и различни. Определящи при анализа на координацията са следните критерии:

- Аргументите на координацията съгласуват своите граматични функции, изразени чрез стойностите на атрибутите: **ARG-ST**⁴⁶, **VALENCE**, **MOD** и **SLASH**.
- Аргументите могат да имат различни синтактични категории (напр. AP и AdvP) и в този случай цялата координация получава само общ етикет за координация.

На базата на тези два критерия съществуват следните типове координация: *наситена координация* (saturated coordination), *адюнктна координация* (adjunct coordination) и *ненаситена координация* (unsaturated coordination). Класификацията според синтактичната категория на фразата - NP координация, VP координация и подобни, играе второстепенна роля. Наситената координация се реализира при празни валентни списъци, атрибут **MOD** със стойност *none* и атрибут **SLASH** с празен списък като стойност. Адюнктната координация е възможна при празни валентни списъци и атрибут **MOD** със стойност, различна от *none*, и съвместяван от всички аргументи. Ненаситената координация има поне един непразен валентен списък или атрибутът **SLASH** има непразен списък като стойност. Непразните списъци при ненаситените координации се съвместяват от всички аргументи. В повечето случаи аргументите на координацията съвместяват и други характеристики, включително своя етикет, но има

⁴⁴ В граматичната литература нееднократно са дискутирани неудобствата, свързани с определянето на съюза или координиращите се елементи като кандидати за опора. Затова ние се придържаме към концепцията в (Pollard and Sag 1994) за разглеждане на този тип структури като фрази без опора.

⁴⁵ Съюзи, запетай и други препинателни знаци, наричани още координиращи връзки.

⁴⁶ Играе роля при координация на лексикално ниво.

примери, когато това не е така. Тъй като категорията не може да липсва, имаме нужда от специална стойност за този случай на несъвместяване. За целта променяме йерархията от типове, като въвеждаме типа *coordination* на същото ниво, на което се намират типовете *noun*, *verb*, *adj*, *prep* и т.н. Следва коригираната йерархия на типовете:

```

head
substantive (subst)
  PRD : boolean
  MOD : mod-synsem
  coordination
    noun
      CASE : case
    verb
      VFORM : vform
      AUX : boolean
      INV : boolean
    adj
    adv
    prep
      PFORM : pform

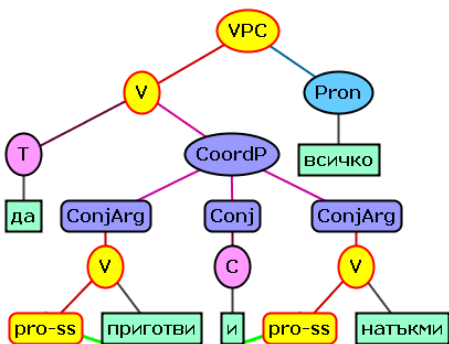
```

Така имаме следните случаи: (1) ако синтактичните категории на аргументите на координацията са съвместими, то координацията получава съответната категория; (2) ако те не са съвместими, то координацията получава категория *coordination*. Тази промяна в йерархията на типовете влияе и на дефинициите за селективната способност на лексикалните знакове в лексикона. Основно следствие е, че координационна фраза с категория *coordination* не може да бъде избрана от друг знак, т.е. не може да е комплемент, подлог или опора на адюнкт. Така получаваме следните видове координация:

Лексикална координация

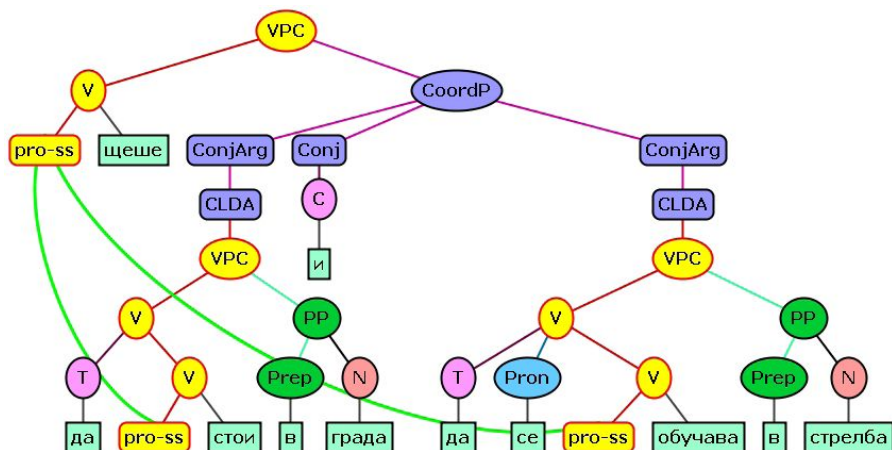
Обикновено всички лексикални знакове могат да участват в координация на лексикално ниво (с изключение на клитиките). Резултатът е, че координацията е лексикален знак, който има същите стойности за валентните списъци и атрибута **MOD**, както и лексикалните знаци на аргументите. Следва пример:

Тук глаголната лексикална координация е в рамките на 'да-конструкция'. Координацията в този случай е лексикален знак, който задължително наследява стойността на атрибута **ARG-ST** от двата глагола.

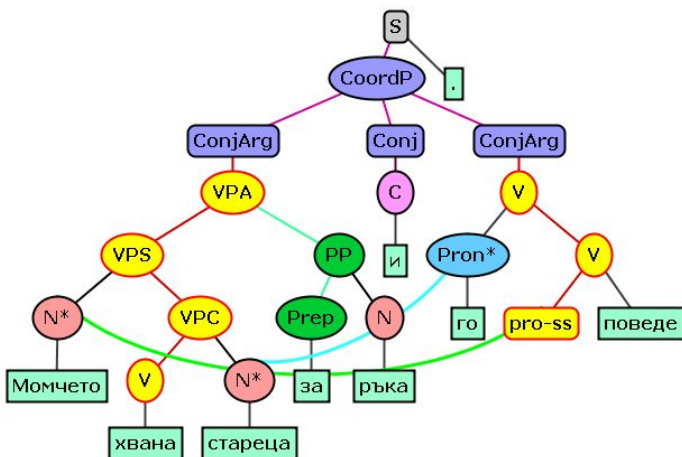


Координация на ниво изречение

Модалните и някои от спомагателните глаголи избират подчинени изречения като компленти. В тези случаи координацията не е проблем, тъй като подчинените изречения са наситени глаголни фрази и следователно техните валентни списъци са празни. Примерът, който следва, показва координацията на две подчинени изречения. При тях прилагането на анализа на горната лексикална структура не е възможно по принцип, тъй като съответните стойности на двата атрибута **ARG-ST** са несъвместими. Координацията се избира от спомагателния глагол *щеше*. Обърнете внимание, че неизразените подлози на трите глагола са свързани с отношение на кореферентност, което означава, че имат един и същ подлог:

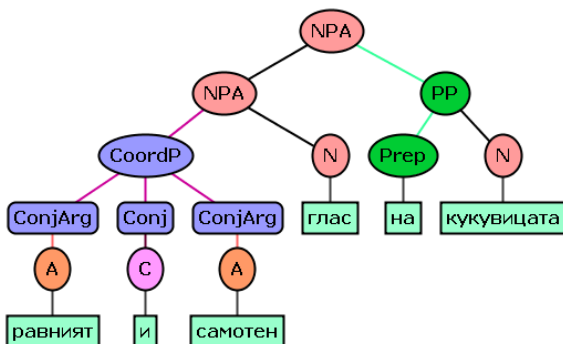


Следващият пример показва координацията на изреченско ниво. Двата глагола имат еднаква валентност, но те не могат да се координират на по-ниско ниво, тъй като към първата глаголна фраза има адюнкт за *ръка*, който се реализира след подлога. Второто изречение има неизразен подлог, който се кореферира с изразения подлог на първото изречение. Има корефериращо отношение и между комплента на първото изречение и винителната клитика във второто.



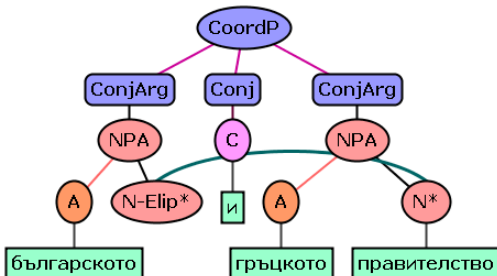
Адюнктна координация в рамките на именната фраза

Адюнктната координацията в рамките на именната фраза е координация на адюнкти на името опора – прилагателни, наречия, относителни подчинени изречения и предложни фрази. Пример:



Тук координацията е между две прилагателни, които модифицират опората **глас** и имат една и съща стойност на своите атрибути **MOD**.

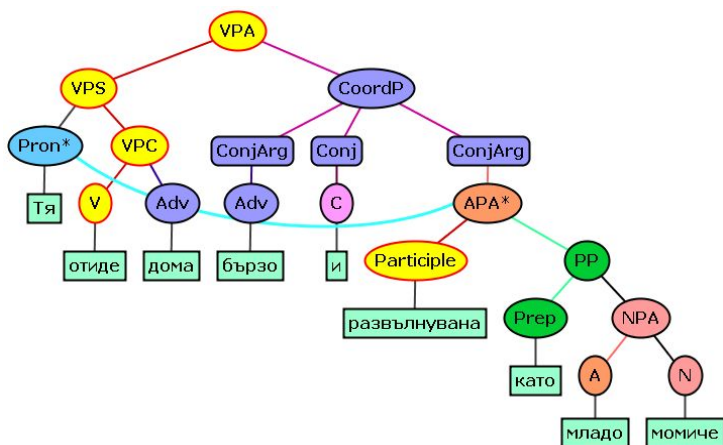
Понякога обаче подобна координация не е възможна, тъй като двете прилагателни не се съгласуват по отношение на стойностите си на атрибута **MOD**:



Потенциалната координация *българското и гръцкото* не е възможна, тъй като стойностите на атрибута **MOD** за прилагателното *българското* и на атрибута **MOD** на прилагателното *гръцкото* не могат да се съвместят. В подобни случаи използваме елипса на съществителното (или именна фраза). Следователно има координация на две именни фрази.

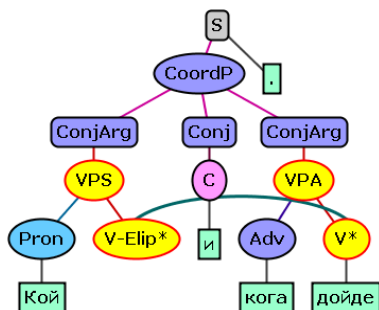
Аджунктна координация в глаголната фраза

Аджункти в рамките на глаголните фрази могат да образуват координация, ако стойностите на техните атрибути **MOD** се съвместяват. В този случай категорията на координацията не играе роля и следователно може да се използва по-общата категория *coordination* без загуба на информация. В примера има координация на два аджункта – единият е наречие, което модифицира цялото *VPS*. Вторият е адективна фраза, която играе роля на вторична предикация и се свързва само с подлога. Следователно двете стойности на атрибутите **MOD** на двата аджункта не си противоречат и могат да бъдат съвместени. Фактът, че вторият аджункт е в ролята на вторична предикация, е отразен в примера чрез кореференция с подлога.



Относно координацията в някои въпросителни изречения

Понякога селективната способност на опората може да бъде нарушена от свойствата на въпросителните думи. Например това се отнася до възможната координация на подлог и аджункт. В този случай нашият анализ включва глаголна елипса, за да има правилно реализиране на зависимите части. При такава елипса координацията става на нивото на изречението:



Елиптичните конструкции представляват много често явление в езика и играят важна роля при дефинирането на формална граматика за даден език. Ние изхождаме от дефиницията: „*Елиптичната конструкция* е конструкция, в която липсват един или няколко елемента, които са възстановими или предвидими от контекста.”⁴⁷ Липсващия елемент ще наричаме *елипса*, а елементите в контекста, които спомагат за възстановяването на липсващия елемент, ще наричаме *носители на елипсата*. Така дефинирана, елиптичната конструкция покрива голямо многообразие от пропуснати части. Това са неизразеният подлог или комплемент, неизразени части на координационните фрази, липсващи предикати и/или комплементи в подчинени изречения, дискурсно⁴⁸ възстановими липсващи елементи. Следователно елипсата може да се лицензира от граматиката на езика (неизразен подлог), от локалния или нелокалния контекст (изпускане на предикат или аргумент, който не е подлог) или от екстралингвистични фактори.

Беше споменато, че субстантивацията и номинализацията са явления, които се лицензират от елипсата на опора или модификатор. Поведението им обаче изисква по-задълбочено проучване.

Тук разглеждаме елипси, които се възстановяват на базата на копиране на характеристики от носителите им. Тези носители на елипси могат да бъдат конкретни елементи в текста или да са част от нашето знание. Понякога има вариативност при възможностите за възстановяване на елипсите. За да маркираме типа носител на елипса, въвеждаме явно представяне на липсващите елементи с помощта на т.нар. *елипсни маркери*. Те показват конституентната позиция на липсващия елемент и имат връзка към елементите, от които се копира липсващата информация. Според нас елипсните маркери показват словоредната позиция на липсващия елемент и неговата конституентна структура. Така ние допускаме липсващите елементи да бъдат не просто една дума, а цяла фраза със съответната структура. Елипсите се различават от извлечените фрази, при които словоредната позиция невинаги може да бъде определена.

Следва представяне на двата типа елипсни маркери, които въвеждаме в нашата граматика.

⁴⁷ An elliptical construction is a construction that lacks an element that is recoverable or inferable from the context (WhatIsAnEllipticalConstruction.htm)

⁴⁸ Имаме предвид възстановяване в по-широкия контекст (напр. в диалог или интервю) или извън текста – от знанията ни за света.

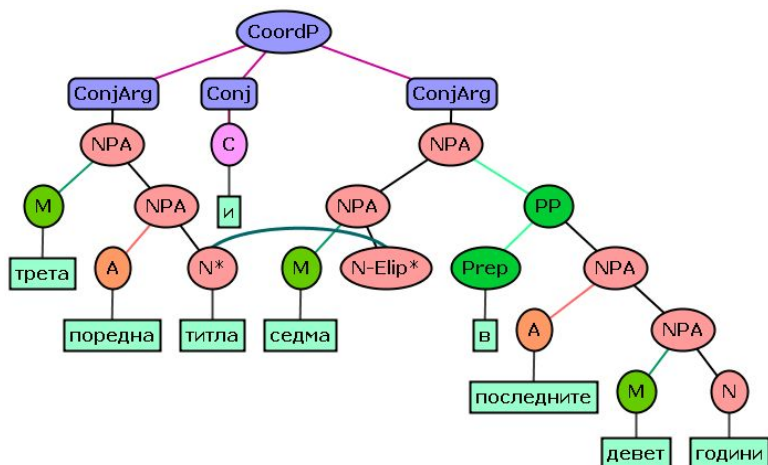
Елипси с елипсни носители в рамките на изречението

Ние разглеждаме елиптични маркери за класовете пълнозначни думи (като включваме и предлозите) и техните области на действие – *V-Elip* (за глаголи и глаголни фрази), *N-Elip* (за съществителни и именни фрази), *A-Elip* (за прилагателни и адективни фрази), *Adv-Elip* (за наречия и адвербиални фрази), *G-Elip* (за деепричастия), *Prep-Elip* (за предлози) и *PP-Elip* (за предложни фрази). Видът на елипсата зависи от вида на носителя на елипсата. Ако носителът е лексикален знак, то елипсата е също лексикален знак. Ако е фраза, то елипсата е фраза от този вид и със същата структура. Следователно елипсният маркер се свързва с максималната фраза на носителя си. Трябва да се има предвид, че при елипсите се копира само формата на носителя (синтактична структура, семантика), но контекстът се променя. Това означава, че съответните елементи на структурата се свързват с нови референти. Ако има кореферентни отношения между носителя на дадена елипса и елементите на възстановената фраза, то те не следват от копирането на формата на носителя, а от други фактори при интерпретацията на изречението, след възстановяването на липсващия елемент.

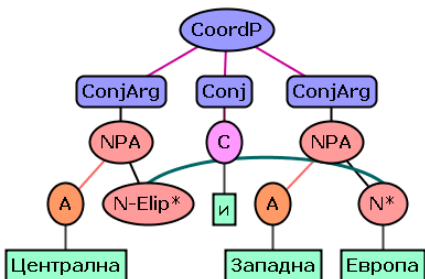
При това възстановяване допускаме модификация на граматично ниво. Например в изречението *Иван мие колата си и Гошо също възстановяването е възможно по два начина: Иван мие колата си и Гошо също мие колата си и Иван мие колата си и Гошо също мие колата му*. При първото възстановяване има нужда само от един елипсен маркер *V-Elip*, който да се свърже с *VPC 'мие колата си'*. Така интерпретацията на изречението е, че всеки един от двамата мие своята собствена кола. Във втория случай се нуждаем от два елипсни маркера – един за глагола и един за съществителното – *V-Elip*, който се свързва с *'мие'*, и *N-Elip*, който се свързва с *'колата си'*, но с промяна на клитиката от *'си'* на *'му'*. В този случай интерпретацията на изречението е, че двамата мият колата на Иван.

Управлението на тези промени при копирането на съответната форма се осъществява с помощта на атрибути, дефинирани за дадения вид елипсен маркер. Например за *V-Elip* това са атрибутите *type* и *gram*. Първият определя три възможности: копиране без промени на словоформата, копиране с промени на словоформата, копиране с отрицателната форма на словоформата. Вторият определя детайлните граматични характеристики на копираната глаголна фраза. Необходимо е копирането с отрицание да бъде отделно, тъй като то определя в някои случаи промяна на лексикалния елемент (вж. по-долу). Някои елипсни маркери нямат атрибути – например, *Adv-Elip*, *G-Elip*, *Prep-Elip* и *PP-Elip*. Елипсният маркер наследява аргументната структура и/или валентните списъци на носителя си. По този начин той може да бъде опора или зависима част. В частност *V-Elip* може да има неизразен подлог. Следват примери на такива елипси.

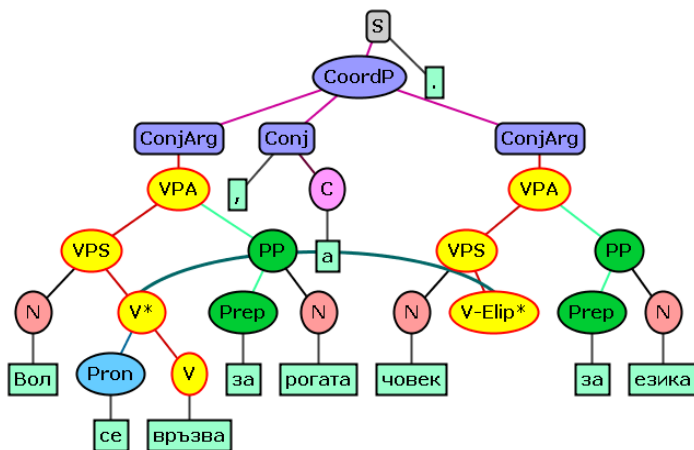
По-долу е показана именна елипса, при която се използва елипсният маркер *N-Elip*. Съществителното *титла* се копира от първия във втория аргумент на координацията.



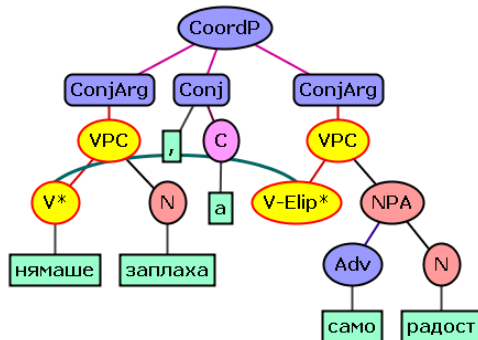
Както е споменато в раздела, посветен на координацията, много често при фраза от вида: **A1 и A2 N**, където A1 и A2 са прилагателни и N е съществително, се блокира анализът на координацията между двете прилагателни, които после да поясняват опората: (A1 + A2) N. Това се получава, тъй като двете прилагателни имат несъвместими изисквания към своята опора и следователно те не могат да се свържат с един и същ референтен индекс. В този случай има елипса на съществителното в рамките на първия аргумент на координацията. Следва един пример на такава елипса: *Централна и Западна Европа*. Тъй като двете части на Европа са с различни референти, въвеждаме елипса, която дублира името Европа:



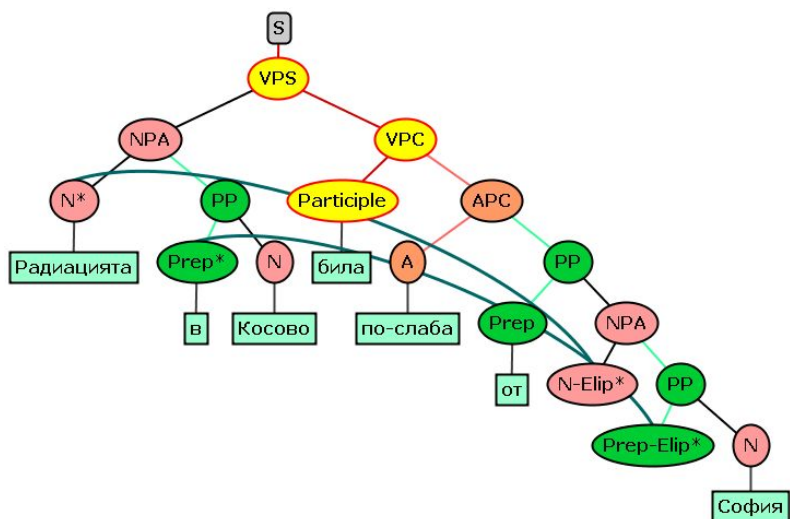
В следващия пример се използва маркерът *V-Elip*, който посочва мястото, където се копира глаголът носител без никакви промени:



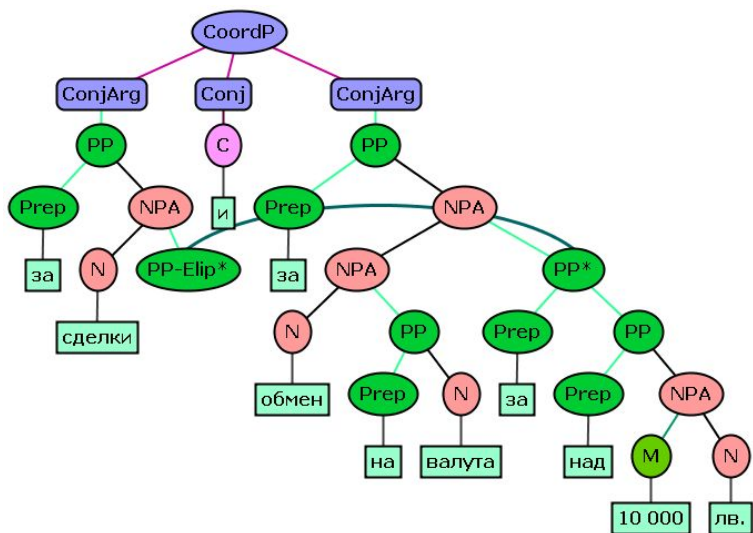
Долният пример показва как маркерът *V-Elip* копира от носителя информация с лексикално отрицание. Примерът е на базата на изречението: *Нямаше заплаха, а само радост*. Тук глаголът 'нямаше' се копира като глагола 'имаше'.



Следва пример, който показва елипса на два елемента. Елипсните маркери *Prep-Elip* и *N-Elip* образуват последователно предложна и именна фраза. Предложният маркер *Prep-Elip* се свързва със съществителното 'София' и образува предложна фраза, която е адюнкт на именния елипсен маркер *N-Elip*.

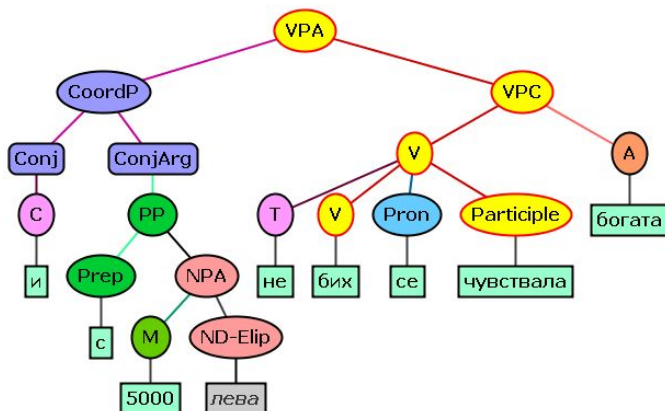


Последният пример показва елипсата на цяла предложна фраза.



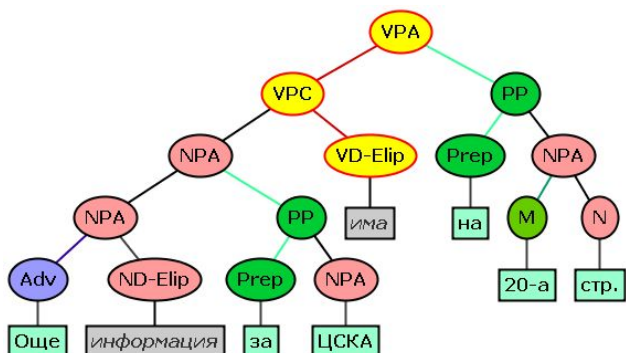
Елипси с елипсни носители извън изречението

Това са случаи, при които носителят на елипсата е в друго изречение от текста (случай, често срещан при диалози), подразбира се от контекста или от общото знание за света. Въвеждаме три елипсни маркера за този вид елипси: *VD-Elip*, *ND-Elip* и *PPD-Elip*. В нашия корпус не сме попадали на други видове елипси от този тип, но принципно граматиката не ги изключва. Тъй като няма явен носител на елипсите в този случай, всички елипсни маркери имат следните три атрибута: *type*, *gram* и *form*. Атрибутът *type* може да взема две стойности: *world_knowledge* и *discourse*. Стойността *world_knowledge* означава, че липсващият елемент може да бъде възстановен само на базата на нашето знание за света. Стойността *discourse* означава, че носителят на елипсата се намира в рамките на текста. Маркерът *VD-Elip* има допълнителна стойност *exists*, която определя елипсата като изпускане на копулата. Това е един много чест случай на елипса от този вид. Останалите атрибути имат стойности като на другите маркери за елипси, описани по-горе. Следват примери:

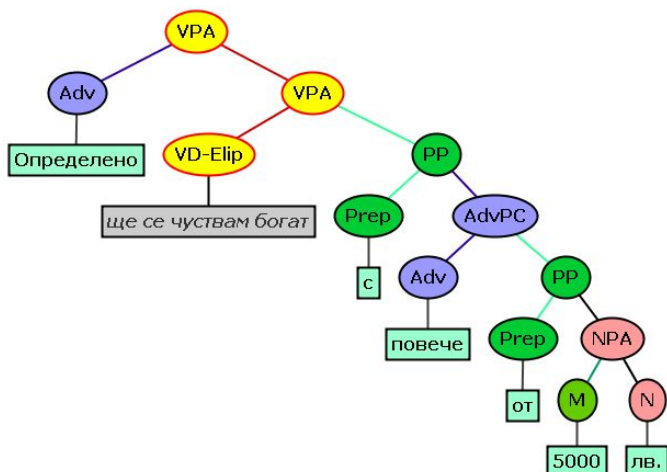


Това е изречение с именна елипса (маркер *ND-Elip*), която се определя от дискурса (в случая – интервю с въпрос „Бихте ли се чувствали богата в България с 5000 лв.“). Словоформата под елипсията е стойността на атрибута *form*.

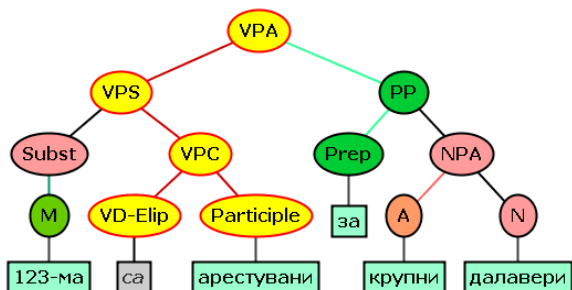
Следващият пример показва два елипсни маркера *ND-Elip* и *VD-Elip*. Именната елипса се възстановява като *‘информация’* на базата на нашето знание за света. Глаголната елипса е от тип *‘exists’* с форма *‘има/намира се’*. Нашето знание в случая се отнася до структурата на един вестник.



Следващият пример демонстрира възстановяването на цяла фраза на базата на дискурса (отново интервю):



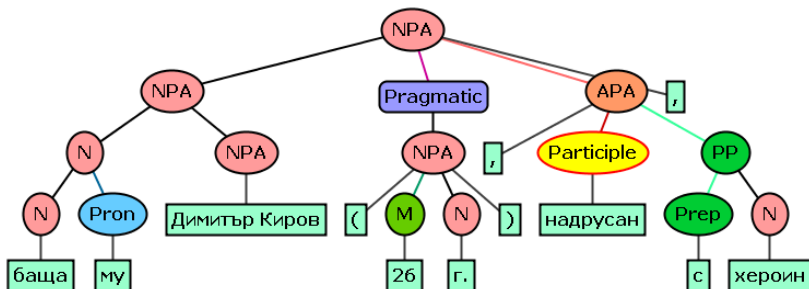
Последният пример е един типичен случай на *VD-Elip* в заглавие на статия от вестник, където копулата системно се изпуска:



Под *прагматични фрази* разбираме фрази, които имат принос към интерпретацията на дадена фраза на прагматично равнище. Анализираме прагматичните фрази като адюнкти на съответните опори (именни, глаголни и т.н.). За да представим коректно този вид адюнкти, разширяваме йерархията от типове за конституентната структура, като въвеждаме два нови подтипа на типа опора-адюнкт – *опора-семантичен адюнкт* и *опора-прагматичен адюнкт*. При графичното им изобразяване не въвеждаме отделни възли за всеки подтип, а ги изобразяваме като адюнктен възел. Прагматичната фраза се реализира между зависимите части на опората. За да се различава от семантичните адюнкти, тя се маркира с функционалния елемент *Pragmatic*. Най-честите прагматични фрази са парентетичните изрази, оценъчните наречия и звателните форми. За детайлен анализ на звателните форми вж. (Osenova and Simov 2002) и по-нататък в раздела. Представеният там анализ може да бъде използван и за анализ на другите прагматични фрази с модификация при съпоставянето на стойността на атрибута **MOD** на прагматичната фраза със синтактично-семантичните характеристики на опората.

Парентетични изрази

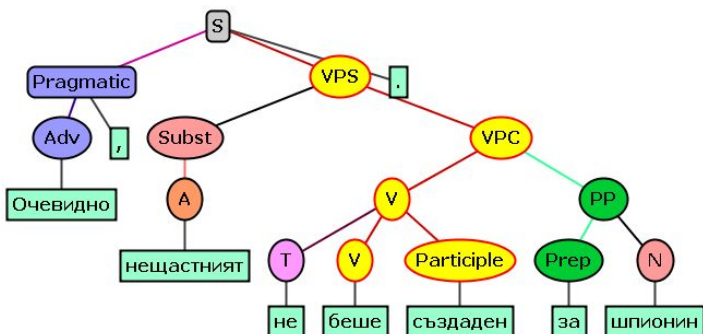
Тук отнасяме изрази, които внасят допълнителна информация към прагматичния контекст на фразата, в рамката на която се появяват, или изразяват отношение на говорещия към съдържанието на изказването. Типични примери за фрази, които изразяват отношението на говорещия, са: *разбира се, от една страна, от друга страна, изглежда, предполагам, за радост, за съжаление, честно казано* и други. Парентетичните изрази, които изразяват допълнителна информация, са по-широк клас. Типични примери са фрази, уточняващи: характеристики на даден човек като години, произход и подобни – *Иван Петров (17 г.)*, Жером Голмар (**Фр**); местоположение на град – *Ванкувър (Канада)*; обстоятелства на действие – *президентът, при посещението си в Пловдив, изказа отрицателно мнение по този въпрос* и други. Следва пример:



Тук елементът *Pragmatic* предоставя информация за възрастта на човека, обозначен от именната фраза.

Оценъчни наречия

Това са наречия, които изразяват модалност или характеристики на цялото изказване. Тези наречия се свързват към корена на дървото, съответстващо на цялото изречение. Ето един пример:



Някои от най-честите такива наречия в корпуса са: *например, може би, сигурно, вероятно, просто, всъщност, очевидно, именно, все пак, явно*. Проблем възниква, когато едно и също наречие може да изпълнява прагматична и непрагматична роля. В такива случаи решаващ е контекстът и/или интерпретацията. Напр. думата *изобщо* играе подобна двойна роля. Сrv. *Защо изобщо трябва да чета?* и *Ти изобщо не ме слушаш*. В първия пример наречието е прагматично и е синонимно на израза *по принцип*. Второто изречение обаче разкрива функцията му като наречие за количество и степен. В този случай то е синонимно на наречието *съвсем*.

Звателни форми

Този вид прагматични фрази ще бъде разгледан по-подробно с оглед на формализирането му в ОФГ. Анализът е представителен и за другите типове прагматични фрази.

Традиционно българските звателни форми се разглеждат спрямо опозицията: звателна форма (остатък от падежната система на имената) – основна форма на името, т.е. с оглед на наличието или неналичието на специални флексии. Но от функционална гледна точка звателните форми включват не само структурната звателна форма от парадигмата на някои имена (съществителните в единствено число от мъжки и женски род, както и дългите форми на някои прилагателни). Всъщност звателността може да се изрази по няколко начина: със звателни форми на имената, с имена в основна форма и посредством различни частици.

Ето класификацията, представена у (Иванова и Ницолова 1995, стр. 24-29):

1. звателни частици, които не назовават адресата, а само насочват към него – *хей, бе*
2. звателни форми, които назовават адресата
 - a. служат само за установяване на контакт
 - b. квалифицират адресата.

Има мнение, че в последния случай има скрита пропозиция с перформативния глагол 'смятам, че някой притежава определено качество' (Ницолова 1984, стр. 44).

В българската граматична литература на звателните форми се приписват основно три функции според синтактичната им позиция в изречението:

- На обикновени обръщения, т.е. парентетични елементи, които не се съгласуват със сказуемото по лице и число и следователно не могат да се интерпретират като подлози (Попов 1998, стр. 130; Брезински 2000, стр. 94):

Господине, днес ще вали.

- На подлози, когато са използвани във въпросителни, възклицателни и подбудителни изречения и се съгласуват със сказуемото по лице и число. Освен това има изискване да не са дублирани от местоимение във второ лице (Попов 1998, стр. 129; Брезински 2000, стр. 94; Граматика на съвременния български книжовен език 1983, стр. 120):

Господине, елате насам!

- На приложения, когато са дублирани от местоимение във второ лице (Попов 1998, стр. 130; Ницолова 1984, стр. 43)

Вие, господине, елате насам!

Според нас съществува унифициращ механизъм, който адекватно показва синтактико-прагматичното поведение на звателните форми, без да разделя и противопоставя различните им употреби. Нашите лингвистични аргументи за подобно становище са следните:

- Ако звателните форми функционират единствено като маркери на дискурса и нямат отношение към изречението, то как да си обясним възможността им за кореферентно отношение с различни части на изречението (с подлога, допълнението, притежателните местоимения и т.н.)
- Ако приемем, че в българския език се приписва субектна роля само на дума в именителен падеж, тогава е необосновано да се избира контекст, в който на звателните форми се приписва такава роля. Съгласуването на подлога с глагола не е достатъчно силен аргумент, тъй като съществуват случаи, при които: (1) между експлицитния подлог и сказуемото няма съгласуване 'Всички излязохме на разходка'; (2) съгласуването е колеба-

ещо се между отделните съставки на подлога ‘Част от студентите влезе/влязоха’ или (3) то е зависимо от семантиката на конюнктите при координация ‘Радост и тъга лъхаше от душата му’. Вж. по-подробно за първите два типа (Osenova 2001).

- Звателните форми се отнасят по един и същи начин към изреченията с различни типове илокуционна сила, включително и декларативните:

Господине, изпуснахте влака.

- Както беше споменато, българският език се причислява към нулевосубектните (pro-drop) езици. В случаите, когато става въпрос за звателни употреби от типа ‘хей ти’, можем да приемем, че местоимението във второ лице притежава две синкретични форми – една в именителен падеж и една в звателна форма.
- Звателните форми са извън областта на лявата дислокация и лявата периферия, които в българския език са свързани с удвояването на допълненията (Пенчев 1993, стр. 120) и с комплементизаторите (Krapova and Karastaneva 1999).

Той Иван го няма.

Нашето предложение е двата възгледа за звателните форми да се комбинират в един, т.е. да ги интерпретираме едновременно като адюнкти и дискурсни маркери.

Те са адюнкти, защото много често споделят синтактични характеристики с различни части в изречението. Но те са и дискурсни маркери, защото взаимодействието със самото изречение се осъществява на надизреченско ниво.

Друг факт, който подкрепя предложената тук идея за звателната форма като адюнкт, е сравнително свободният словоред на именните звателни форми. Срв. например възможните позиционни варианти на звателната форма ‘господине’ в следните изречения:

Господине, елате насам!

Елате, господине, насам!

Елате насам, господине!

Един възможен контрааргумент може да бъде, че остенсивните частици⁴⁹, местоименията и по-сложните звателни фрази са ограничени в дистрибуцията си на обръщения, но според нас този факт по никакъв начин не противоречи на тезата ни. Всички тези средства за изразяване се вписват в адюнктната интерпретация, тъй като типично за адюнктите е да са разнотипни и рекурсивни.

⁴⁹ Тук използваме думата ‘остенсивен’ в значението на ‘остенсивен стимул’. Това е понятие в Теорията на релевантността и означава начин да се привлече вниманието на слушателя към казаното от говорещия.

В резултат на нашия модуларен поглед върху явлението предлагаме следните видове отношения между синтактичната и прагматичната специфика на звателните форми:

- звателната форма и изразеният/неизразеният подлог или допълнението на изречението са кореферентни, т.е. посочват към един и същ обект в действителността.

Господине, (вие) изпуснахте влака.

Господине, търсят Ви.

- звателната форма контролира притежателно или рефлексивно местоимение в изречението.

Жена Ви, господине, се обаждат.

Господине, спрете жена си!

- звателните форми служат като остенсивен стимул по отношение на слушателя относно някакъв факт.

Господине, навън вали.

Преди да преминем към самото формализиране на звателните форми, ще отделим място на характеристиките **CONTENT** и **CONTEXT**, тъй като те са особено релевантни за правилното им анализиране.

Стойността на характеристиката **CONTENT** може да бъде една от следните: *psoa* (вм. parametrized-state-of-affairs) *параметрично състояние на света*, *nom-obj* (вм. nominal-object) *именен обект* и *quant* (вм. quantifier) *квантификатор*. Първият тип описва семантиката на глаголите, прилагателните, предложните фрази и имената в предикативна употреба. Вторият тип описва семантиката на именните фрази и обикновено се представя като **RESTR** (*ограничение*) и **INDEX** (*индекс*). Индексът се разделя на два подвида: *ref* (референтен) и *non-ref* (нереферентен). Ограничението задава характеристиките на обекта, който се реферира чрез именната фраза, а индексът⁵⁰ служи за свързването с обект от действителността и се представя чрез съгласувателните характеристики: род, число и лице. Третият тип *quant* описва семантиката на квантификаторите.

Стойността на характеристиката **CONTEXT** има две характеристики: **BACKGROUND**, която описва с помощта на множество от *psoa* ситуацията на речевия акт, **CONTEXTUAL-INDICES** (*контекстни индекси*), които включват описание на **SPEAKER** (*говорещ*), **ADDRESSEE** (*слушател*), **UTTERANCE-LOCATION** (*място на изказване*) и други. Частта от йерархията на типовете за стойността на характеристиката **CONTEXT** е, както следва:

⁵⁰ В ситуационната семантика тази характеристика отговаря на т.нар. дискурсен маркер (discourse marker). Подобни индекси използваме при кореференцията (вж. по-долу).

context

BACKGROUND : *set(psoa)*

CONTEXTUAL-INDICES : *c-inds*

c-inds

SPEAKER : *ref*

ADDRESSEE : *ref*

UTTERANCE-LOCATION : *ref*

На синтактично ниво разглеждаме звателните форми като специален вид адюнкти, които се свързват с изречението чрез тяхната **MOD** характеристика.

Главната причина да изберем подобен подход е, че чрез посочения механизъм изречението става 'видимо' за звателната форма. А това на свой ред е необходимо най-вече в случаите, когато звателната форма споделя различни синтактични характеристики с вътрешноизреченските елементи.

Информацията, която представя звателната форма като название на адресата с определени характеристики, се разпределя в атрибуtnата структура по следния начин:

- присъства като стойност на характеристиката **ADDRESSEE** в рамките на характеристиката **CONTEXT** на изречението, като по този начин осигурява факта, че звателната форма съвпада със слушателя на това изречение. Освен това при свързването си със слушателя в **ADDRESSEE** звателната форма се коиндексира по съответния начин с частите на изречението;
- присъства в рамките на характеристиката **CONTENT** на звателната форма;
- присъства в характеристиката **ADDRESSEE** в рамките на характеристиката **CONTEXT** на звателната форма, където обединява ограниченията на референтно свързване (**RESTR**), идващи от изречението (като опорна дъщеря) и звателната форма (като адюнктна дъщеря);
- присъства в характеристиката **CONTEXT** на най-високото ниво, където всъщност съвпада с кодираната в **CONTEXT|ADDR** информация на звателната форма.

Един проблем, който възниква при интерпретирането на звателните форми като адюнкти, е Семантичният принцип в ОФГ (вж. по-подробно в раздела за Опорната фразова граматика). Нека припомним, че този принцип изисква характеристиката **CONTENT** на най-високия възел да е същата като характеристиката **CONTENT** на семантичната опора. А в конструкции от типа *опора-адюнкт* семантичната опора е адюнктът. Затова в нашия анализ на звателните форми променяме това положение, уточнявайки, че звателната форма не може да бъде семантична опора.

За да осъществим идеята си, се налага да въведем специален принцип, тъй като в (Pollard and Sag 1994, p. 337) не са формулирани принципи, които да действат върху характеристиката **C-INDICES** на дъщерите на дадена фраза. В същото време сме наясно, че има разработки в рамките на характеристиката **CONTEXT** с цел разпознаване на диалогови фрагменти в Опорната фразова граматика (вж. Ginzburg and Lappin 2001). Но за сегашните ни цели, а именно – да анализираме първо отношението между звателните форми и изреченията, ние предлагаме следния принцип:

Принцип за звателните форми:

*Когато в една структура от типа опора-адюнкт, адюнктната дъщеря е в звателна форма, то стойността на характеристиката **ADDRESSEE** на звателната форма съвпада със стойността на характеристиката **ADDRESSEE** на най-високия възел.*

Освен въвеждането на горния принцип, ние променяме и дефиницията за семантична опора. Тази промяна може да се смята за резултат от прагматичната същност на звателните форми и да се възприеме като преместване на информационната релевантност от семантиката към прагматиката, т.е. от характеристиката **CONTENT** в характеристиката **CONTEXT**.

Когато в една структура от типа опора-адюнкт адюнктната дъщеря е в звателна форма, то опорната дъщеря става семантична опора. Във всички останали случаи семантичната опора е адюнктът.

Сега вече е възможно да представим общия формален механизъм на взаимодействие между звателната форма и изречението:

- Цялото изречение има стойност на характеристиката **DTRS** от типа *опора-адюнкт*. Компонентите в рамките на този тип са следните: опорната дъщеря е глаголна фраза с наситена валентност (saturated), а адюнктната дъщеря е фраза в звателна форма.
- Начинът, по който звателната форма ‘избира’ своята опора, е кодиран в рамките на самата нея. Така стойността на **MOD** характеристиката на адюнкта изисква за опора глаголна фраза с наситена валентност, т.е. изречение.
- Стойността на **MOD** характеристиката на звателната форма е същата като стойността на характеристиката **SYNSEM** на опорната дъщеря и по този начин стойността на характеристиката **ADDRESSEE** на опорната дъщеря става достъпна в структурата на адюнктната дъщеря. Така стойността на характеристиката **INDEX** в рамките на характеристиката **ADDRESSEE (ADDR)** на звателната форма е коиндексирана със стойността на характеристиката **INDEX** в рамките на характеристиката **CONTENT** на звателната форма, а също така

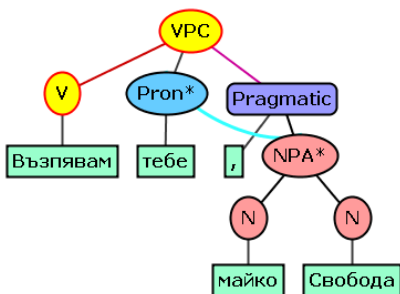
и със стойността на характеристиката **INDEX** в рамките на характеристиката **ADDRESSEE** на опорната дъщеря. Ограничението (**RESTR**) на **ADDRESSEE** на звателната фраза е обединение от ограниченията в **CONTENT** на звателната форма и ограниченията на **ADDRESSEE** на опорната дъщеря.

- Посредством Принципа за звателните форми, дефиниран по-горе, стойността на характеристиката **ADDRESSEE** на звателната форма е същата като стойността на характеристиката **ADDRESSEE** на максималната фраза.

Сега нека представим по-общо идеята за формалните механизми, които осигуряват взаимодействието между синтаксис и прагматика в рамките на трите звателни типа, които вече бяха дискутирани.

- Звателната форма и изразеният/неизразеният подлог или допълнението в изречението се отнасят към един и същ обект от действителността. В рамките на опорната дъщеря **CONTENT** характеристиката на подлога (или съответно на допълнението) се свързва по съответния начин с характеристиката **ADDRESSEE** в рамките на **CONTEXT**.
- Звателната форма контролира притежателно или рефлексивно местоимение в изречението. Този механизъм се извършва първо в опорното изречение. Там притежателното или рефлексивното местоимение се контролира от подлога на изречението. Характеристиката **CONTENT** на подлога се коиндексира по подходящ начин с характеристиката **ADDRESSEE** на изречението. Оттук се коиндексира със звателната форма.
- Звателната форма служи просто като остенсивен стимул за слушателя относно някакъв факт. В този случай отново има подходящо коиндексиране между стойността на характеристиката **ADDRESSEE** и характеристиката **CONTENT** на звателната форма, но характеристиката **INDEX** в **ADDRESSEE** не съвпада с никой **INDEX** от характеристиката **CONTENT** на изречението.

За визуализация на звателната форма в дървовидната структура използваме функционалния възел *Pragmatic*. Също така се отбелязва и кореференцията между звателната фраза и съответните елементи на изречението, ако има такава. Следва един пример:



Тук звателната фраза *майко Свобода* е свързана с елемент от тип *VPC*, но модифицира прагматичните характеристики на цялото изречение. Звателната фраза е в кореферентна релация с винителното местоимение във второ лице, което е комплемент във *VPC* фразата.

Приемаме, че всеки знак е свързан с поне един **семантичен индекс**. Този индекс може да обозначава обект или множество от обекти в света. Именните фрази и някои видове подчинени изречения въвеждат поне един индекс в областта на интерпретиране на тези фрази чрез своята семантика – стойността на атрибута **CONTENT**. Прилагателните, наречията, адективните и адвербиалните фрази, адюнктните подчинени изречения изискват съществуването на подходящ индекс, с който да се свържат и да модифицират опората с помощта на атрибута **MOD**. В зависимост от семантиката си всяка фраза може да въвежда повече от един индекс. Например притежателните местоимения освен индекса, достъпен чрез атрибута **MOD** (който се свързва с индекса на притежавания обект, качество и т.н.), въвеждат и втори индекс за притежателя; относителните прилагателни, освен индекса, управляван от атрибута **MOD**, въвеждат и втори индекс за обекта, който е в съответното отношение с първия индекс (*дървена поставка* – един индекс за поставката и един за дървото, от което е направена); глаголите, освен индекса за събитието, което въвеждат, имат индекси за своите задължителни аргументи и за аргументите си по премълчаване (дефинирани локално в речниковата статия или наследени от йерархизираната структура на речника); аналогично за съществителни, въвеждащи или предполагащи събития. Метонимните употреби също въвеждат повече от един индекс.

Това многообразие от индекси и знание, свързано с тях, излиза извън обсега на една формална граматика. За да бъде представено и формализирано задоволително, то изисква наличието на формални средства като *онтология на света* и *конкретни знания за състоянието на света*. Тук е описана само една част от релациите, които следват от това знание за света и които играят важна роля за правилния синтактичен анализ на много конструкции в езика. **Кореферентно отношение** е всяка релация между два или повече индекса. Така дефинирани, тези релации покриват огромно количество отношения, повечето от които се съдържат в онтологията на света. Като най-често срещани въвеждаме и използваме следните три отношения:

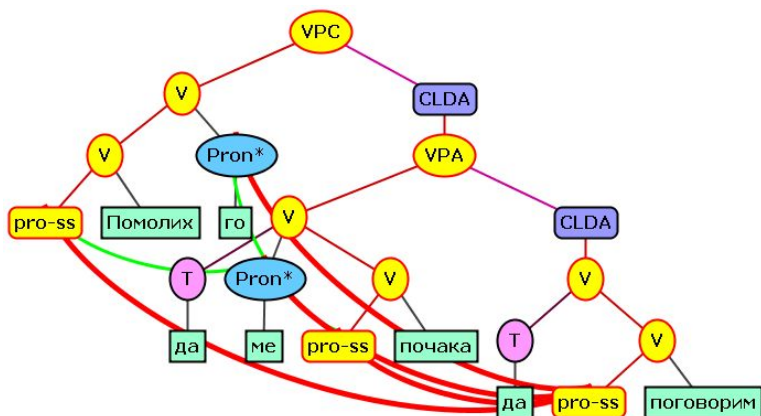
equal(x,y) – равенство на два индекса;

member-of(x,y) – *x* е елемент на множеството *y*;

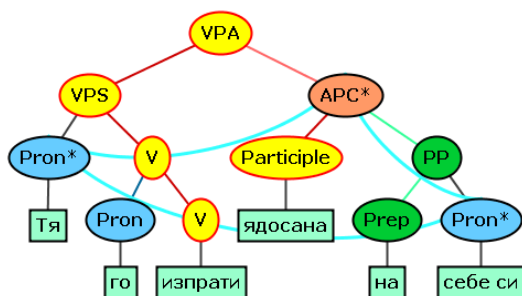
subset-of(x,y) – множеството *x* е подмножество на множеството *y*.

Тези релации са достатъчни за анализирането на едно голямо количество от явления като: анафорични отношения; интерпретация на неизразени аргументи (най-често подлог); съгласуване; вторична предикация; определяне на при-

тежателя при притежателни местоимения; номинализация; кохерентни вериги в текста и други. Самите илюстративни примери в текста на книгата представят различни кореферентни отношения. Затова тук даваме само два допълнителни примера. Първият показва използване на отношението *member-of(x,y)*, а вторият е свързан с анализа на вторичната предикация.



В горния пример имаме две отношения на равенство. Едното е между индекса на неизразения подлог на глагола *помолих* и винителната кратка форма на личното местоимение в първо лице *ме*. Другото е между индекса на неизразения подлог на глагола *почака* и винителната кратка форма на личното местоимение в трето лице *го*. Имаме и четири отношения *member-of(x,y)* между тези четири индекса и индекса, въведен от неизразения подлог на глагола *поговорим*.



Вторичната предикация винаги се изразява с помощта на адюнкт към съответната глаголна фраза. В горния пример индексът на адюнктната APC фраза се кореферира с индекса на подлога и индекса на възвратноличното местоимение. Трите отношения са равенство.

Формализирането на граматиката на даден естествен език представлява прилагане на определена теоретична рамка върху широк спектър от езиков материал. Ние избрахме Опорната фразова граматика за наша отправна точка по няколко причини. Най-напред, тя разполага със солиден формалистичен апарат, който я прави подходяща за компютърни приложения. След това, тя разглежда нивата на езика (фонетика, морфология, лексикология, синтаксис, семантика, прагматика) в тяхната взаимосвързаност. Това увеличава значително обяснителната сила при анализите и позволява да се интерпретират посложни и проблемни езикови явления. Не на последно място, ОФГ балансира между теориите, които работят с фразови структури (конституенти) и теориите, които описват зависимостта между думите (депенденти).

Идеите, представени в настоящата книга, стоят в основата на анотационната схема, по която бяха анализирани морфологично и синтактично изреченията в корпуса BulTreeBank (15000 изречения). Схемата беше приложена върху изолирани и свързани изречения от различни жанрове, но с преобладаващо съдържание от съвременните периодични издания и съвременната художествена литература. Принципите, от които се ръководехме, бяха непротиворечивост и модуларност при анализите, както и изчерпателност при обхвата на явленията.

В рамките на тази книга беше невъзможно да се разгледат в голяма детайлност езиковите нива и връзките между тях. Затова предпочетохме да наблегнем на пълнотата на типовете езикови елементи пред детайлността на описанията. Приоритет беше даден също така на спорните и интересни случаи. При описанието на явленията решихме да представим такива, които нямат еднозначно решение в езиковата теория (координация, елипса, прагматични елементи) и такива, които присъстват неотменно в езиковия анализ (кореференцията). Разбира се, остават неописани още редица не по-малко интересни явления като: ролята на детерминацията на междумфразово и изреченско ниво, вторичната предикация и др., които ще бъдат обект на бъдеща разработка.

Създаването на корпус от синтактични описания на българския език има принос не само за теоретичното преосмисляне и допълване на езиковите факти. Корпусът е ценен ресурс за различни приложни цели. От гледна точка на лингвистиката такъв архив може да се използва за създаване на речници, за подпомагане на обучението и превода, за тестване на езикови хипотези. От гледна точка на информационните технологии наличието на добре анализиран корпус предполага създаването на автоматични средства за компютърна обработка на българския език. Макар че е необходима още много работа за осъществяването на автоматичен анализ на български изречения – което в крайна сметка е една от нашите главни цели, постигнатото досега представлява една солидна основа за разширяване на лингвистичните знания и корпуси, от една страна, и построяване на софтуерни анализиращи програми, от друга страна.

Използвана литература

Алмалех 1993: Мони Алмалех. Семантика и синтаксис. Университетско издателство „Св. Кл. Охридски“, София.

Байрамова 1982: Мая Байрамова. Рецензия: A. Grannes. Loan Compounds in Bulgarian Reflecting the Turkish Indefinite Izafet-Construction. Oslo – Bergen – Troms, Universitetforlaget, 1980. 56 стр. В: Съпоставително езикознание, 1982, VII, № 4, стр. 58-67.

Бояджиев, Куцаров и Пенчев 1998: Тодор Бояджиев, Иван Куцаров и Йордан Пенчев. Съвременен български език. Издателска къща „Петър Берон“, София.

Брезински 2000: Стефан Брезински. Български синтаксис. София.

Граматика на съвременния български книжовен език. Синтаксис. 1983, Издателство на БАН, София.

Иванова и Ницолова 1995: Калина Иванова и Руселина Ницолова. Ние, говорещите хора. София.

Куцаров 2007: Иван Куцаров. Теоретична граматика на българския език. Морфология. Университетско издателство „Паисий Хилендарски“, Пловдив.

Ницолова 1984: Руселина Ницолова. Прагматичен аспект на изречението в българския книжовен език. София.

Ницолова 1986: Руселина Ницолова. Българските местоимения. София.

Осенова 2004: Петя Осенова. Именните фрази от типа NN. В: Сборник от Шестите национални славистични четения: Славистиката в началото на XXI век. Традиции и очаквания. София, 26-27.04.2002 г.

Осенова 2007: Петя Осенова. Въпросът за опората и депendentите в българските атрибутивни именни групи от гледна точка на Опорната фразова граматика (ОФГ). В: Opera Slavica, 4, Praha.

Паскалева 2007: Елена Паскалева. Компютърна морфология. София.

Пенчев 1962: Йордан Пенчев. Мястото на определения-прилагателни при местоименията-съществителни нещо и нищо. В: Известия на ИБЕ, кн. VIII.

Пенчев 1993: Йордан Пенчев. Синтаксис. Управление и свързване. Пловдив.

Попов 1998: Константин Попов. Синтаксис на българския книжовен език. Велико Търново.

Симов и Колковска 2004: Кирил Симов и Сия Колковска. Интерпретация на дак-конструкциите в Опорната фразова граматика. В: Сборник от Шестите национални славистични четения: Славистиката в началото на XXI век. Традиции и очаквания. София, 26-27.04.2002 г.

Тишева 2005: Йовка Тишева. За системата на общите въпроси в съвременния български език. В: Аргументна структура. Сема РШ, София, стр. 160-182.

Butt et al. 1999: Miriam Butt, Tracy H. King, Maria-Eugenia Nino and Frederique Segond. *A Grammar Writer's Cookbook*. CSLI, Stanford.

Ginzburg and Lappin 2001: Gregory J. Ginzburg and Sh. Lappin SHARDS: Fragment Resolution in Dialogue. In: Bunt H., van der Slijs, Thijsse eds., *Proc. of the 4th International conference on Computational Semantics*, Tilburg, pp. 156-172.

Krapova and Karastaneva 1999: I. Krapova and T. Karastaneva. On the structure of The CP field in Bulgarian. In: *FASSL – 1999, Working Papers in Linguistics*, 20, ed. by M. Vulchanova and L. Hellan, Trondheim

Manning and Sag 1999: Christopher D. Manning and Ivan A. Sag. Dissociations between Argument Structure and Grammatical Relations. In: Webelhuth, Gert, Jean-Pierre Koenig and Andreas Kathol (eds.), *Lexical And Constructional Aspects of Linguistic Explanation*, CSLI Publications, pp. 63-78.

Osenova 2001: Petya Osenova. On Subject-Verb Agreement in Bulgarian (An HPSG-based Account). In: *Proceedings of the Fourth FDSL*, Potsdam, pp. 661-672.

Osenova 2007: Petya Osenova. A Corpus-based Approach To Free Relatives with „който” in Bulgarian. In: *Proceedings from 5th Conference on FASSBL*, 18-20 Sept. 2006, pp. 109-112.

Osenova and Simov 2002: Petya Osenova and Kiril Simov. Bulgarian Vocative within HPSG Framework. In: *Proceedings of the 9th International Conference on HPSG*, J.-B. Kim, S. Wechsler (eds.), CSLI Publications, pp. 233-244.

Polard and Sag 1994: Carl Pollard and Ivan Sag. *Head-Driven Phrase Structure Grammar*. CSLI Publications.

Pustejovsky 1995: James Pustejovsky. *The Generative Lexicon*. The MIT Press.

Sag et. al 1993: Ivan Sag, Tom Wasow and Emily Bender. *Syntactic Theory*. Stanford

Simov, Osenova and Slavcheva 2004: Kiril Simov, Petya Osenova and Milena Slavcheva. BTB-TR03: BulTreeBank Morphosyntactic Tagset. BulTreeBank Project Technical Report № 03. At: <http://www.bultreebank.org/TechRep.html>

Tisheva and Dzhonova 2002: Yovka Tisheva and Marina Dzhonova. Information Structure Level in Treebanks. In: *Proceedings of The First Workshop on Treebanks and Linguistic Theories*. 20-21 September, 2002, Sozopol, pp. 231-252.

What Is An Elliptical Construction.htm.

At: <http://www.sil.org/linguistics/glossaryoflinguisticterms/>, 15 Sept. 2007

Приложение 1

Списък на синтактичните етикети

(по азбучен ред)

A	Лексикален елемент <i>прилагателно име</i>
A-Elip	Елипса на прилагателно или адекватна фраза, които се възстановяват в рамките на изречението
Adv	Лексикален елемент <i>наречие</i>
Adv-Elip	Елипса на наречие или адвербиална фраза, които се възстановяват в рамките на изречението
APA	Фраза на прилагателното от типа <i>опора – адюнкт</i>
APC	Фраза на прилагателното от типа <i>опора – комплемент</i>
AdvPA	Фраза на наречието от типа <i>опора – адюнкт</i>
AdvPC	Фраза на наречието от типа <i>опора – комплемент</i>
C	Лексикален елемент <i>съюз</i>
CL	Подчинено изречение
CLCHE	Подчинено изречение, въведено със съюза <i>че</i>
CLDA	Подчинено изречение с <i>да-конструкция</i>
CLQ	Подчинено изречение, въведено с <i>въпросително местоимение</i>
CLR	Подчинено изречение, въведено с <i>относително местоимение</i>
CLZADA	Подчинено изречение, въведено със съюза <i>за да</i>
Conj	<i>Координационен съюз</i> в рамките на координационна структура
CoordP	<i>Координационна структура</i>
ConjArg	<i>Координиращите се елементи</i> в рамките на координационната структура
DiscA	Зависим елемент, чиято словоредна реализация е между опората, която го управлява, и по-нисък елемент
DiscE	Зависим елемент, чиято словоредна реализация е по-високо от опората, която го управлява
DiscM	Зависимите елементи на две фрази, които са смесени, без никоя от тях да управлява другата
G-Elip	Елипсично деепричастие, което се възстановява в рамките на изречението
Gerund	Лексикален елемент <i>деепричастие</i>
H	Лексикален елемент <i>фамилно име</i> или <i>прилагателно, образувано от собствено име</i>
I	Лексикален елемент <i>междуметие</i>
M	Лексикален елемент <i>числително име</i>
N	Лексикален елемент <i>съществително име</i>

N-Elip	Елипса на съществително или именна фраза, които се възстановяват в рамките на изречението
ND-Elip	Елипса на съществително или именна фраза, които се възстановяват в рамките на надизреченския контекст
nid	Зависим елемент, изразен извън обсега на фразата
Nomin	Предикат, междуметие или друга не-именна фраза във функция на именен елемент
NPA	Именна фраза от типа <i>опора – адюнкт</i>
NPC	Именна фраза от типа <i>опора – комплемент</i>
PP	Предложна фраза от типа <i>опора – адюнкт</i> или <i>опора – комплемент</i>
PP-Elip	Елипса на предложна фраза, която се възстановява в рамките на изречението
PPD-Elip	Елипса на предложна фраза, която се възстановява в рамките на надизреченския контекст
Pragmatic	Прагматичен елемент
Prep	Лексикален елемент <i>предлог</i>
Prep-Elip	Елипса на предлог, който се възстановява в рамките на изречението
Participle	Лексикален елемент <i>причастие</i>
Pron	Лексикален елемент <i>местоимение</i>
pro-ss	Маркер за неизразен подлог
Subst	Зависим елемент (прилагателно, числително, причастие или местоимение), който при елипса на именната опора поема функцията ѝ
T	Лексикален елемент <i>частица</i>
V	Лексикален елемент <i>глагол</i>
V-Elip	Елипса на глагол или глаголна фраза, които се възстановяват в рамките на изречението
VD-Elip	Елипса на глагол или глаголна фраза, които се възстановяват в рамките на надизреченския контекст
Verbalized	Местоимение, частица или наречие във функция на предикат
VPA	Глаголна фраза от типа <i>опора – адюнкт</i>
VPC	Глаголна фраза от типа <i>опора – комплемент</i>
VPF	Глаголна фраза от типа <i>опора – запълваща част</i>
VPS	Глаголна фраза от типа <i>опора – подлог</i>
+f	Маркер за фраза с опора в <i>първа словоредна позиция</i>
+l	Маркер за фраза с опора в <i>последна словоредна позиция</i>

Приложение 2

Списък на морфологично-синтактичните етикети, дефинирани и използвани в проекта BulTreeBank

(по части на речта)

Морфологично-синтактичните етикети представляват последователност от букви, цифри и тирета, където всяка позиция отговаря на една граматична характеристика. Първата буква е винаги главна и отговаря на частта на речта. Ако дадена характеристика не е съвместима с даден член на парадигмата, тя се замества с тире. За детайлно описание на множеството от морфологично-синтактични етикети, вж. (Simov, Osenova, Slavcheva 2004). Тук даваме само списък на етикетите за справка. Всеки етикет съответства на член на парадигмата на даден клас думи. Някои етикети не са дадени, с цел пестене на място, но могат да се възстановят на базата на други етикети.

Съществително име – N			
Ncmssi	сщ.нар., м.р., ед.ч., нчл.	Npmssi	сщ.собст., м.р., ед.ч., нчл.
Ncmsh	сщ.нар., м.р., ед.ч., кр.ч.	Npmsh	сщ.собст., м.р., ед.ч., кр.ч.
Ncmssf	сщ.нар., м.р., ед.ч., пл.ч.	Npmssf	сщ.собст., м.р., ед.ч., пл.ч.
Ncmstd	сщ.нар., м.р., ед.ч., чл.	Npmsv	сщ.собст., м.р., ед.ч., зв.ф.
Ncmsv	сщ.нар., м.р., ед.ч., зв.ф.	Npmsa	сщ.собст., м.р., ед.ч., в.п.
Ncmsa	сщ.нар., м.р., ед.ч., в.п.	Npmsd	сщ.собст., м.р., ед.ч., д.п.
Ncmsd	сщ.нар., м.р., ед.ч., д.п.	Npmppi	сщ.собст., м.р., мн.ч., нчл.
Ncmppi	сщ.нар., м.р., мн.ч., нчл.	Npmpd	сщ.собст., м.р., мн.ч., чл.
Ncmppd	сщ.нар., м.р., мн.ч., чл.	Npmt	сщ.собст., м.р., бр.ф.
Ncmt	сщ.нар., м.р., бр.ф.	Npfsi	сщ.собст., ж.р., ед.ч., нчл.
Ncfssi	сщ.нар., ж.р., ед.ч., нчл.	Npfstd	сщ.собст., ж.р., ед.ч., чл.
Ncfstd	сщ.нар., ж.р., ед.ч., чл.	Npfpi	сщ.собст., ж.р., мн.ч., нчл.
Ncfpi	сщ.нар., ж.р., мн.ч., нчл.	Npfpd	сщ.собст., ж.р., мн.ч., чл.
Ncfpd	сщ.нар., ж.р., мн.ч., чл.	Npfsv	сщ.собст., ж.р., ед.ч., зв.ф.
Ncfsv	сщ.нар., ж.р., ед.ч., зв.ф.	Npnsi	сщ.собст., ср.р., ед.ч., нчл.
Ncnsi	сщ.нар., ср.р., ед.ч., нчл.	Npnstd	сщ.собст., ср.р., ед.ч., чл.
Ncnstd	сщ.нар., ср.р., ед.ч., чл.	Npnpi	сщ.собст., ср.р., мн.ч., нчл.
Ncnpi	сщ.нар., ср.р., мн.ч., нчл.	Npnpd	сщ.собст., ср.р., мн.ч., чл.
Ncnpd	сщ.нар., ср.р., мн.ч., чл.	Np-li	сщ.собст., мн.ч., пл.т., нчл.
Nc-li	сщ.нар., мн.ч., пл.т., нчл.	Np-lid	сщ.собст., мн.ч., пл.т., чл.
Nc-lid	сщ.нар., мн.ч., пл.т., чл.		

Прилагателно име – A			
Amsi	м.р., ед.ч., нчл.	Afsi	ж.р., ед.ч., нчл.
Amsh	м.р., ед.ч., кр.ч.	Afstd	ж.р., ед.ч., чл.
Amsf	м.р., ед.ч., пл.ч.	Ansi	ср.р., ед.ч., нчл.
Ams-e	м.р., ед.ч., д.ф.	Anstd	ср.р., ед.ч., чл.
Amsa	м.р., ед.ч., в.п.	A-pi	мн.ч., нчл.
Amsd	м.р., ед.ч., д.п.	A-pd	мн.ч., чл.

Фамилни или прилагателни имена – Н

Hmsi	м.р., ед.ч., нчл.	Hnsi	ср.р., ед.ч., нчл.
Hmsh	м.р., ед.ч., кр.ч.	Hnsd	ср.р., ед.ч., чл.
Hmsf	м.р., ед.ч., пл.ч.	H-pi	мн.ч., нчл.
Hfsi	ж.р., ед.ч., нчл.	H-pd	мн.ч., чл.
Hfsd	ж.р., ед.ч., чл.		

Лични местоимения – Рр

Ppe-os1	им.п., ед.ч., 1 л.	Ppetds3m	кр.ф., д.п., ед.ч., 3 л., м.р.
Ppe-os2	им.п., ед.ч., 2 л.	Ppelds3n	д.п., ед.ч., 3 л., ср.р.
Ppe-os3m	им.п., ед.ч., 3 л., м.р.	Ppetds3n	кр.ф., д.п., ед.ч., 3 л., ср.р.
Ppe-os3f	им.п., ед.ч., 3 л., ж.р.	Ppelds3f	д.п., ед.ч., 3 л., ж.р.
Ppe-os3n	им.п., ед.ч., 3 л., ср.р.	Ppetds3f	кр.ф., д.п., ед.ч., 3 л., ж.р.
Ppe-op1	им.п., мн.ч., 1 л.	Ppeldp1	д.п., мн.ч., 1 л.
Ppe-op2	им.п., мн.ч., 2 л.	Ppetdp1	кр.ф., д.п., мн.ч., 1 л.
Ppe-op3	им.п., мн.ч., 3 л.	Ppeldp2	д.п., мн.ч., 2 л.
Ppelas1	в.п., ед.ч., 1 л.	Ppetdp2	кр.ф., д.п., мн.ч., 2 л.
Ppetas1	кр.ф., в.п., ед.ч., 1 л.	Ppeldp3	д.п., мн.ч., 3 л.
Ppelas2	в.п., ед.ч., 2 л.	Ppetdp3	кр.ф., д.п., мн.ч., 3 л.
Ppetas2	кр.ф., в.п., ед.ч., 2 л.	Ppetss1	кр.ф., д.п.пр., ед.ч., 1 л.
Ppelas3m	в.п., ед.ч., 3 л., м.р.	Ppetss2	кр.ф., д.п.пр., ед.ч., 2 л.
Ppetas3m	кр.ф., в.п., ед.ч., 3 л., м.р.	Ppetss3m	кр.ф., д.п.пр., ед.ч., 3 л., м.р.
Ppelas3f	в.п., ед.ч., 3 л., ж.р.	Ppetss3f	кр.ф., д.п.пр., ед.ч., 3 л., ж.р.
Ppetas3f	кр.ф., в.п., ед.ч., 3 л., ж.р.	Ppetss3n	кр.ф., д.п.пр., ед.ч., 3 л., ср.р.
Ppelas3n	в.п., ед.ч., 3 л., ср.р.	Ppetsp1	кр.ф., д.п.пр., мн.ч., 1 л.
Ppetas3n	кр.ф., в.п., ед.ч., 3 л., ср.р.	Ppetsp2	кр.ф., д.п.пр., мн.ч., 2 л.
Ppelap1	в.п., мн.ч., 1 л.	Ppetsp3	кр.ф., д.п.пр., мн.ч., 3 л.
Ppetap1	кр.ф., в.п., мн.ч., 1 л.	Pph-os2	л., уч.ф., им.п., ед.ч., 2 л.
Ppetap2	в.п., мн.ч., 2 л.	Pphlas2	л., уч.ф., в.п., ед.ч., 2 л.
Ppetap2	кр.ф., в.п., мн.ч., 2 л.	Pphtas2	л., уч.ф., кр.ф., в.п., ед.ч., 2 л.
Ppelap3	в.п., мн.ч., 3 л.	Pphlds2	л., уч.ф., д.п., ед.ч., 2 л.
Ppetap3	кр.ф., в.п., мн.ч., 3 л.	Pphtds2	л., уч.ф., кр.ф., д.п., ед.ч., 2 л.
Ppelds1	д.п., ед.ч., 1 л.	Pphtss2	л., уч.ф., кр.ф., д.п.пр., ед.ч., 2 л.
Ppetds1	кр.ф., д.п., ед.ч., 1 л.	Ppxta	възвр.л., кр.ф., в.п.
Ppelds2	д.п., ед.ч., 2 л.	Ppxla	възвр.л., в.п.
Ppetds2	кр.ф., д.п., ед.ч., 2 л.	Ppxtd	възвр.л., кр.ф., д.п.
Ppelds3m	д.п., ед.ч., 3 л., м.р.	Ppxts	възвр.л., кр.ф., д.п.пр.

Показателни местоимения – Pd

Pde-os-m	им.п., ед.ч., м.р.	Pda--s-f	за призн., ед.ч., ж.р.
Pde-os-f	им.п., ед.ч., ж.р.	Pda--s-n	за призн., ед.ч., ср.р.
Pde-os-n	им.п., ед.ч., ср.р.	Pda--p	за призн., мн.ч.
Pde-op	им.п., мн.ч.	Pdl	за място
Pde-as-m	в.п., ед.ч., м.р.	Pdm	за начин
Pde-ds-m	д.п., ед.ч., м.р.	Pdq	за кол.
Pda--s-m	за призн., ед.ч., м.р.	Pdt	за време

Относителни местоимения – Pr			
Pre-os-m	им.п., ед.ч., м.р.	Pra--p	за призн., мн.ч.
Pre-os-f	им.п., ед.ч., ж.р.	Prl	за място
Pre-os-n	им.п., ед.ч., ср.р.	Prm	за начин
Pre-op	им.п., мн.ч.	Prp--s-m	за притеж., ед.ч., м.р.
Pre-as-m	в.п., ед.ч., м.р.	Prp--s-f	за притеж., ед.ч., ж.р.
Pre-ds-m	д.п., ед.ч., м.р.	Prp--s-n	за притеж., ед.ч., ср.р.
Pre--s	ед.ч.	Prp--p	за притеж., мн.ч.
Pra--s-m	за призн., ед.ч., м.р.	Prq	за кол.
Pra--s-f	за призн., ед.ч., ж.р.	Prt	за време
Pra--s-n	за призн., ед.ч., ср.р.		

Обобщителни местоимения – Pc			
Pce-os-m	им.п., ед.ч., м.р.	Pca--p	за призн., мн.ч.
Pce-os-f	им.п., ед.ч., ж.р.	Pcl	за място
Pce-os-n	им.п., ед.ч., ср.р.	Pcm	за начин
Pce-op	им.п., мн.ч.	Pcq--s-mh	за кол., ед.ч., м.р., кр.ч.
Pce-as-m	в.п., ед.ч., м.р.	Pcq--s-mf	за кол., ед.ч., м.р., пл.ч.
Pce-ds-m	д.п., ед.ч., м.р.	Pcq--s-fd	за кол., ед.ч., ж.р., чл.
Pca--s-m	за призн., ед.ч., м.р.	Pcq--s-nd	за кол., ед.ч., ср.р., чл.
Pca--s-f	за призн., ед.ч., ж.р.	Pcq--p--d	за кол., мн.ч., чл.
Pca--s-n	за призн., ед.ч., ср.р.	Pct	за време

Въпросителни местоимения – Pi			
Pie-os-m	им.п., ед.ч., м.р.	Pil	за място
Pie-os-f	им.п., ед.ч., ж.р.	Pim	за начин
Pie-os-n	им.п., ед.ч., ср.р.	Pic	за причина
Pie-op	им.п., мн.ч.	Pip--s-m	за притеж., ед.ч., м.р.
Pie-as-m	в.п., ед.ч., м.р.	Pip--s-f	за притеж., ед.ч., ж.р.
Pie-ds-m	д.п., ед.ч., м.р.	Pip--s-n	за притеж., ед.ч., ср.р.
Pia--s-m	за призн., ед.ч., м.р.	Pip--p	за притеж., мн.ч.
Pia--s-f	за призн., ед.ч., ж.р.	Piq	за кол.
Pia--s-n	за призн., ед.ч., ср.р.	Piy	за брой лица
Pia--p	за призн., мн.ч.	Pit	за време

Неопределителни местоимения – Pf			
Pfe-os-mi	им.п., ед.ч., м.р., нчл.	Pfe-os-f	им.п., ед.ч., ж.р.
Pfe-os-mh	им.п., ед.ч., м.р., кр.ч.	Pfe-os-n	им.п., ед.ч., ср.р.
Pfe-os-mf	им.п., ед.ч., м.р., пл.ч.	Pfe-op	им.п., мн.ч.
Pfe-os-fi	им.п., ед.ч., ж.р., нчл.	Pfe-as-m	в.п., ед.ч., м.р.
Pfe-os-fd	им.п., ед.ч., ж.р., чл.	Pfe-ds-m	д.п., ед.ч., м.р.
Pfe-os-ni	им.п., ед.ч., ср.р., нчл.	Pfe--s-n	ед.ч., ср.р.
Pfe-os-nd	им.п., ед.ч., ср.р., чл.	Pfa--s-m	за призн., ед.ч., м.р.
Pfe-op--i	им.п., мн.ч., нчл.	Pfa--s-f	за призн., ед.ч., ж.р.
Pfe-op--d	им.п., мн.ч., чл.	Pfa--s-n	за призн., ед.ч., ср.р.
Pfe-os-m	им.п., ед.ч., м.р.	Pfa--p	за призн., мн.ч.

Pfl	за място	Pfq	за кол.
Pfm	за начин	Pfq-----i	за кол., нчл.
Pfp--s-m	за притеж., ед.ч., м.р.	Pfq-----d	за кол., чл.
Pfp--s-f	за притеж., ед.ч., ж.р.	Pfy-----i	за брой лица, нчл.
Pfp--s-n	за притеж., ед.ч., ср.р.	Pfy-----d	за брой лица, чл.
Pfp--p	за притеж., мн.ч.	Pft	за време

Отрицателни местоимения – Pn

Pne-os-m	им.п., ед.ч., м.р.	Pna--s-n	за призн., ед.ч., ср.р.
Pne-os-f	им.п., ед.ч., ж.р.	Pna--p	за призн., мн.ч.
Pne-os-n	им.п., ед.ч., ср.р.	Pnl	за място
Pne-os-ni	им.п., ед.ч., ср.р., нчл.	Pnm	за начин
Pne-os-nd	им.п., ед.ч., ср.р., чл.	Pnp--s-m	за притеж., ед.ч., м.р.
Pne-op	им.п., мн.ч.	Pnp--s-f	за притеж., ед.ч., ж.р.
Pne-as-m	в.п., ед.ч., м.р.	Pnp--s-n	за притеж., ед.ч., ср.р.
Pne-ds-m	д.п., ед.ч., м.р.	Pnp--p	за притеж., мн.ч.
Pna--s-m	за призн., ед.ч., м.р.	Pnq	за кол.
Pna--s-f	за призн., ед.ч., ж.р.	Pnt	за време

Притежателни местоимения – Ps

един притежател

Psol-s1mi	ед.ч., 1 л., м.р., нчл.	Psol-s3fim	ед.ч., 3 л., ж.р., нчл., м.р.прит.
Psol-s1mh	ед.ч., 1 л., м.р., кр.ч.	Psol-s3fdm	ед.ч., 3 л., ж.р., чл., м.р.прит.
Psol-s1mf	ед.ч., 1 л., м.р., пл.ч.	Psol-s3nim	ед.ч., 3 л., ср.р., нчл., м.р.прит.
Psol-s1fi	ед.ч., 1 л., ж.р., нчл.	Psol-s3ndm	ед.ч., 3 л., ср.р., чл., м.р.прит.
Psol-s1fd	ед.ч., 1 л., ж.р., чл.	Psol-p3-im	мн.ч., 3 л., нчл., м.р.прит.
Psol-s1ni	ед.ч., 1 л., ср.р., нчл.	Psol-p3-dm	мн.ч., 3 л., чл., м.р.прит.
Psol-s1nd	ед.ч., 1 л., ср.р., чл.	Psot--3--m	кр.ф., 3 л., м.р.прит.
Psol-p1-i	мн.ч., 1 л., нчл.	Psol-s3mif	ед.ч., 3 л., м.р., нчл., ж.р.прит.
Psol-p1-d	мн.ч., 1 л., чл.	Psol-s3mhf	ед.ч., 3 л., м.р., кр.ч., ж.р.прит.
Psot--1	кр.ф., 1 л.	Psol-s3mff	ед.ч., 3 л., м.р., пл.ч., ж.р.прит.
Psol-s2mi	ед.ч., 2 л., м.р., нчл.	Psol-s3fif	ед.ч., 3 л., ж.р., нчл., ж.р.прит.
Psol-s2mh	ед.ч., 2 л., м.р., кр.ч.	Psol-s3fdf	ед.ч., 3 л., ж.р., чл., ж.р.прит.
Psol-s2mf	ед.ч., 2 л., м.р., пл.ч.	Psol-s3nif	ед.ч., 3 л., ср.р., нчл., ж.р.прит.
Psol-s2fi	ед.ч., 2 л., ж.р., нчл.	Psol-s3ndf	ед.ч., 3 л., ср.р., чл., ж.р.прит.
Psol-s2fd	ед.ч., 2 л., ж.р., чл.	Psol-p3-if	мн.ч., 3 л., нчл., ж.р.прит.
Psol-s2ni	ед.ч., 2 л., ср.р., нчл.	Psol-p3-df	мн.ч., 3 л., чл., ж.р.прит.
Psol-s2nd	ед.ч., 2 л., ср.р., чл.	Psot--3--f	кр.ф., 3 л., ж.р.прит.
Psol-p2-i	мн.ч., 2 л., нчл.	Psol-s3min	ед.ч., 3 л., м.р., нчл., ср.р.прит.
Psol-p2-d	мн.ч., 2 л., чл.	Psol-s3mhn	ед.ч., 3 л., м.р., кр.ч., ср.р.прит.
Psot--2	кр.ф., 2 л.	Psol-s3mfhn	ед.ч., 3 л., м.р., пл.ч., ср.р.прит.
Psol-s3mim	ед.ч., 3 л., м.р., нчл., м.р.прит.	Psol-s3fin	ед.ч., 3 л., ж.р., нчл., ср.р.прит.
Psol-s3mhmn	ед.ч., 3 л., м.р., кр.ч., м.р.прит.	Psol-s3fdn	ед.ч., 3 л., ж.р., чл., ср.р.прит.
Psol-s3mfmn	ед.ч., 3 л., м.р., пл.ч., м.р.прит.	Psol-s3nin	ед.ч., 3 л., ср.р., нчл., ср.р.прит.
		Psol-s3ndn	ед.ч., 3 л., ср.р., чл., ср.р.прит.
		Psol-p3-in	мн.ч., 3 л., нчл., ср.р.прит.
		Psol-p3-dn	мн.ч., 3 л., чл., ср.р.прит.
		Psot--3--n	кр.ф., 3 л., ср.р.прит.

много притежатели

Pszl-s1mi	ед.ч., 1 л., м.р., нчл.	Pszl-s2ni	ед.ч., 2 л., ср.р., нчл.
Pszl-s1mh	ед.ч., 1 л., м.р., кр.ч.	Pszl-s2nd	ед.ч., 2 л., ср.р., чл.
Pszl-s1mf	ед.ч., 1 л., м.р., пл.ч.	Pszl-p2-i	мн.ч., 2 л., нчл.
Pszl-s1fi	ед.ч., 1 л., ж.р., нчл.	Pszl-p2-d	мн.ч., 2 л., чл.
Pszl-s1fd	ед.ч., 1 л., ж.р., чл.	Pszl--2	кр.ф., 2 л.
Pszl-s1ni	ед.ч., 1 л., ср.р., нчл.	Pszl-s3mi	ед.ч., 3 л., м.р., нчл.
Pszl-s1nd	ед.ч., 1 л., ср.р., чл.	Pszl-s3mh	ед.ч., 3 л., м.р., кр.ч.
Pszl-p1-i	мн.ч., 1 л., нчл.	Pszl-s3mf	ед.ч., 3 л., м.р., пл.ч.
Pszl-p1-d	мн.ч., 1 л., чл.	Pszl-s3fi	ед.ч., 3 л., ж.р., нчл.
Pszl--1	кр.ф., 1 л.	Pszl-s3fd	ед.ч., 3 л., ж.р., чл.
Pszl-s2mi	ед.ч., 2 л., м.р., нчл.	Pszl-s3ni	ед.ч., 3 л., ср.р., нчл.
Pszl-s2mh	ед.ч., 2 л., м.р., кр.ч.	Pszl-s3nd	ед.ч., 3 л., ср.р., чл.
Pszl-s2mf	ед.ч., 2 л., м.р., пл.ч.	Pszl-p3-i	мн.ч., 3 л., нчл.
Pszl-s2fi	ед.ч., 2 л., ж.р., нчл.	Pszl-p3-d	мн.ч., 3 л., чл.
Pszl-s2fd	ед.ч., 2 л., ж.р., чл.	Pszl--3	кр.ф., 3 л.

възвратно

Psxlos-mi	им.п., ед.ч., м.р., нчл.	Psxlos-nd	им.п., ед.ч., ср.р., чл.
Psxlos-mh	им.п., ед.ч., м.р., кр.ч.	Psxlop--i	им.п., мн.ч., нчл.
Psxlos-mf	им.п., ед.ч., м.р., пл.ч.	Psxlop--d	им.п., мн.ч., чл.
Psxlos-fi	им.п., ед.ч., ж.р., нчл.	Psxto	кр.ф., им.п.
Psxlos-fd	им.п., ед.ч., ж.р., чл.	Psxla	в.п.
Psxlos-ni	им.п., ед.ч., ср.р., нчл.	Psxld	д.п.

учтива форма

Pshl-s2mi	ед.ч., 2 л., м.р., нчл.	Pshl-s2ni	ед.ч., 2 л., ср.р., нчл.
Pshl-s2mh	ед.ч., 2 л., м.р., кр.ч.	Pshl-s2nd	ед.ч., 2 л., ср.р., чл.
Pshl-s2mf	ед.ч., 2 л., м.р., пл.ч.	Pshl-p2-i	мн.ч., 2 л., нчл.
Pshl-s2fi	ед.ч., 2 л., ж.р., нчл.	Pshl-p2-d	мн.ч., 2 л., чл.
Pshl-s2fd	ед.ч., 2 л., ж.р., чл.	Psht--2	кр.ф., 2 л.

Числително име – М

Mcmsi	числ., бр., м.р., ед.ч., нчл.	Mc-pd	числ., бр., мн.ч., чл.
Mcmsh	числ., бр., м.р., ед.ч., кр.ч.	Momsi	числ., ред., м.р., ед.ч., нчл.
Mcmsf	числ., бр., м.р., ед.ч., пл.ч.	Momsh	числ., ред., м.р., ед.ч., кр.ч.
Mcmpi	числ., бр., м.р., мн.ч., нчл.	Momshf	числ., ред., м.р., ед.ч., пл.ч.
Mcmpd	числ., бр., м.р., мн.ч., чл.	Mofsi	числ., ред., ж.р., ед.ч., нчл.
Mcfsi	числ., бр., ж.р., ед.ч., нчл.	Mofsd	числ., ред., ж.р., ед.ч., чл.
Mcfsd	числ., бр., ж.р., ед.ч., чл.	Monsi	числ., ред., ср.р., ед.ч., нчл.
Mcfpi	числ., бр., ж.р., мн.ч., нчл.	Monsd	числ., ред., ср.р., ед.ч., чл.
Mcfpd	числ., бр., ж.р., мн.ч., чл.	Mo-pi	числ., ред., мн.ч., нчл.
Mcnsi	числ., бр., ср.р., ед.ч., нчл.	Mo-pd	числ., ред., мн.ч., чл.
Mcnsd	числ., бр., ср.р., ед.ч., чл.	Md-pi	числ., адverb., мн.ч., нчл.
Mcnpi	числ., бр., ср.р., мн.ч., нчл.	Md-pd	числ., адverb., мн.ч., чл.
Mcnpd	числ., бр., ср.р., мн.ч., чл.	My-pi	числ., прибл., мн.ч., нчл.
Mc-pi	числ., бр., мн.ч., нчл.	My-pd	числ., прибл., мн.ч., чл.

Глагол – V

Vpit* *Глагол, личен, несвършен вид, преходен*

Vpitf-r1s	гл., изяв., сег., 1 л., ед.ч.
Vpitf-r2s	гл., изяв., сег., 2 л., ед.ч.
Vpitf-r3s	гл., изяв., сег., 3 л., ед.ч.
Vpitf-r1p	гл., изяв., сег., 1 л., мн.ч.
Vpitf-r2p	гл., изяв., сег., 2 л., мн.ч.
Vpitf-r3p	гл., изяв., сег., 3 л., мн.ч.
Vpitf-o1s	гл., изяв., мсв., 1 л., ед.ч.
Vpitf-o2s	гл., изяв., мсв., 2 л., ед.ч.
Vpitf-o3s	гл., изяв., мсв., 3 л., ед.ч.
Vpitf-o1p	гл., изяв., мсв., 1 л., мн.ч.
Vpitf-o2p	гл., изяв., мсв., 2 л., мн.ч.
Vpitf-o3p	гл., изяв., мсв., 3 л., мн.ч.
Vpitf-m1s	гл., изяв., мнсв., 1 л., ед.ч.
Vpitf-m2s	гл., изяв., мнсв., 2 л., ед.ч.
Vpitf-m3s	гл., изяв., мнсв., 3 л., ед.ч.
Vpitf-m1p	гл., изяв., мнсв., 1 л., мн.ч.
Vpitf-m2p	гл., изяв., мнсв., 2 л., мн.ч.
Vpitf-m3p	гл., изяв., мнсв., 3 л., мн.ч.
Vpitz--2s	гл., пов., 2 л., ед.ч.
Vpitz--2p	гл., пов., 2 л., мн.ч.
Vpitcar-smi	прч., деят., сег., ед.ч., м.р., нчл.
Vpitcar-smh	прч., деят., сег., ед.ч., м.р., кр.ч.
Vpitcar-smf	прч., деят., сег., ед.ч., м.р., пл.ч.
Vpitcar-sfi	прч., деят., сег., ед.ч., ж.р., нчл.
Vpitcar-sfd	прч., деят., сег., ед.ч., ж.р., чл.
Vpitcar-sni	прч., деят., сег., ед.ч., ср.р., нчл.
Vpitcar-snd	прч., деят., сег., ед.ч., ср.р., чл.
Vpitcar-p-i	прч., деят., сег., мн.ч., нчл.
Vpitcar-p-d	прч., деят., сег., мн.ч., чл.
Vpitcao-smi	прч., деят., мсв., ед.ч., м.р., нчл.
Vpitcao-smh	прч., деят., мсв., ед.ч., м.р., кр.ч.
Vpitcao-smf	прч., деят., мсв., ед.ч., м.р., пл.ч.
Vpitcao-sfi	прч., деят., мсв., ед.ч., ж.р., нчл.
Vpitcao-sfd	прч., деят., мсв., ед.ч., ж.р., чл.
Vpitcao-sni	прч., деят., мсв., ед.ч., ср.р., нчл.
Vpitcao-snd	прч., деят., мсв., ед.ч., ср.р., чл.
Vpitcao-p-i	прч., деят., мсв., мн.ч., нчл.
Vpitcao-p-d	прч., деят., мсв., мн.ч., чл.
Vpitcam-smi	прч., деят., мнсв., ед.ч., м.р., нчл.
Vpitcam-sfi	прч., деят., мнсв., ед.ч., ж.р., нчл.
Vpitcam-sni	прч., деят., мнсв., ед.ч., ср.р., нчл.
Vpitcam-p-i	прч., деят., мнсв., мн.ч., нчл.
Vpitcv--smi	прч., страд., ед.ч., м.р., нчл.
Vpitcv--smh	прч., страд., ед.ч., м.р., кр.ч.
Vpitcv--smf	прч., страд., ед.ч., м.р., пл.ч.
Vpitcv--sfi	прч., страд., ед.ч., ж.р., нчл.
Vpitcv--sfd	прч., страд., ед.ч., ж.р., чл.
Vpitcv--sni	прч., страд., ед.ч., ср.р., нчл.
Vpitcv--snd	прч., страд., ед.ч., ср.р., чл.
Vpitcv--p-i	прч., страд., мн.ч., нчл.
Vpitcv--p-d	прч., страд., мн.ч., чл.
Vpitg	деепричастие

Глаголната парадигма се повтаря (с изключение на липсващи форми) за другите типове глаголи, като се променят само първите четири букви на всеки етикет. За икономия тук даваме само съответните префикси. Звездичката означава остатъка от етикета:

Vpii*	гл., лич., несв., непр.
Vppi*	гл., лич., св., непр.
Vnii*	гл., безл., несв., непр.
Vnit*	гл., безл., несв., пр.
Vnpi*	гл., безл., св., непр.
Vnpt*	гл., безл., св., пр.
Vxit*	гл., спом. <i>съм</i> , несв., пр.
Vypt*	гл., спом. <i>бъда</i> , св., пр.
Viit*	гл., спом. <i>бивам</i> , несв., пр.

Наречие – D

Dm	за начин
Dt	за време
DI	за място
Dq	за кол. и степен
Dd	модално

Съюз – C

Cc	съчинителен
Cs	подчинителен
Cr	съотносителен съчинителен
Cr	единичен и съотн. съч.

Частици – T

Ta	утвърдителна
Tn	отрицателна
Ti	въпросителна
Tx	формообразуваща
Tm	модална
Tv	глаголна
Te	за усилване
Tg	за степенуване

Предлог – R

Междуметие – I

ФОРМАЛНА ГРАМАТИКА НА БЪЛГАРСКИЯ ЕЗИК

първо издание

Автори: © Петя Осенова, Кирил Симов

Графичен дизайн: © Николай Генов

Предпечат: Аспектум ООД

ISBN: 978-954-92148-2-6

Издава: Институт за паралелна обработка на информацията, БАН

София, 2007