

Омонимия и синонимия знаков препинания в русском тексте

Т.Ю.Кобзарева

РГГУ, Москва

stamstam@mtu-net.ru

Описываются функции знаков препинания (ЗП) в русском тексте (Т), контекстные особенности актуализации омонимии, синонимии и неоднозначности ЗП, напр., ситуации с запятой: а) оператором сочинения слов, б) границей сегментов (простых или придаточных предложений, обособленных оборотов и др.), в) в сочетании [*и*] сохраняющей морфологическую автономность и значение или теряющей их и т.д. Работа необходима как лингвистический базис любых актуальных задач анализа Т.

1. Введение.

Дешифровка текста при решении задач, предполагающих автоматический анализ, требует, независимо от формы ее реализации, знания возможных интерпретаций всех компонент линейной структуры предложения.

Знаки препинания, по крайней мере в языках с достаточно развитой письменной традицией, образуют специфическую по своей семантике группу. Соответствуя паузам и определенным типам интонирования в устной речи, они и в значительной мере определяют смысл и текста, и предложения. Организуя содержательное членение линейной структуры предложения, ЗП манифестируют соответствие между многомерным графом предложения и его линейной структурой.

Данное рассмотрение не претендует на полноту, но предлагает некоторый способ рассмотрения функций ЗП, значимых для поверхностно-синтаксического анализа¹.

2. Знаки препинания и проективность линейной структуры.

Один из основных аспектов значения ЗП – их функции в предложении.

Сегментация – построение таких компонент предложения (S) как простые-главные и придаточные предложения, деепричастные, обособленные определительные и сравнительные обороты и др. [2] – так или иначе осуществляется в любой модели анализа естественного текста.

Для анализа линейной структуры Т важно, что ЗП могут выполнять две противоположные

функции: 1. связывать слова в проективные отрезки (например, запятая как оператор сочинения слов) внутри сегмента и 2. выступать в качестве операторов членения – границ сегментов, эксплицитно заданных в тексте соответственно правилам русской пунктуации.

При этом одни и те же ЗП могут выступать в Т и в том и в другом качестве.

Для построения проективных отрезков именных, предложных и сочиненных групп, как и для собственно построения сегментов, необходимо знать исходные функции конкретных ЗП и контекстные условия их дешифровки.

Одной из чрезвычайно значимых особенностей именно письменного Т является проективность подчинительных и сочинительных связей [9, 7] внутри сегмента, из которых следует проективность сегмента, т.е. то, что любое слово сегмента, кроме одного – вершины или «скрепы», может находиться в отношениях подчинения и/или сочинения только со словами своего сегмента. Именно проективность сегментов позволяет восстанавливать их целостность при разрывах сегментов вложением в них сегментов-слуг или вводных оборотов [5].

Возможны следующие типы организации линейной структуры со вложенными сегментами: 1. вложенный сегмент находится левее любого слова сегмента-хозяина; 2. вложенный сегмент расположен правее любого слова сегмента-хозяина; 3. вложенный сегмент находится справа от некоторой части сегмента-хозяина и правее него есть хотя бы одно слово сегмента-хозяина.

В последнем случае будем говорить о разрыве сегмента вложением. Многие проблемы сегментации порождаются возможностью разрыва. В каждом конкретном случае важно установить, есть ли разрыв в линейной структуре и, если он есть, найти и соединить разорванные вложением части сегмента-хозяина. Это невозможно без представлений об омонимии и синонимии ЗП, об их

¹ Работа подготовлена при поддержке гранта РФФИ № 03-06-80109. Тема рассмотрена в рамках подхода системы ПСА, разрабатываемой в Отделении интеллектуальных систем РГГУ и представленной в [2].

способности одновременно исполнять несколько функций в предложении и о возможной неоднозначности ЗП, порождающей неоднозначность сегментации предложения.

Таким образом, сегментацию можно рассматривать как процедуру определения значений ЗП в Т.

3. Омонимия и бифункциональность ЗП.

Чтобы определить функции знака в данном Т, нужно иметь исходный набор его возможных значений и знать, как они могут совмещаться. Можно утверждать, что некоторое значение входит в исходный набор значений рассматриваемого ЗП, если существует линейная структура, где это значение у этого ЗП будет единственным.

В случаях, когда ЗП имеет в Т одновременно два значения, т.е. совмещаются две его функции, будем говорить о бифункциональности ЗП.

3.1. Запятая.

Рассмотрим три функции запятой на примере двух S.

Запятая1 – левая граница сегмента. Это значение иллюстрирует ситуация, когда это – левая граница сегмента, вложенного в любой другой сегмент. В S1: *Процесс одевания, которому предается Акакий Акакиевич, шитье и облачение в шинель на самом деле – его разоблачение, постепенный возврат к полной наготе его же призрака.* (Набоков, далее Н) В S1 – три запятые. Только первая из них – запятая1.

Запятая2 – правая граница сегмента. Это значение вычлняется в чистом виде в случае вложения сегмента в сегмент, когда сегмент-хозяин разорван. В предложении S2: *Прохлада монастырской обители, вьющиеся по стенам розы, изможденный человек в ермолке, сияющие, свежие краски фрески, над которой он трудится, – вот та рабочая обстановка, о которой мечтал Гоголь.* (Н) 7 запятых, из них 6-ая запятая (перед тире) представляет чистый случай правой границы.

Запятая3 – оператор сочинения внутри сегмента. В S1 только функцию сочинения слов имеет последняя запятая, манифестирующая сочинение вершин именных групп *его разоблачение и постепенный возврат к полной наготе его же призрака.*

В S2 первые 4 запятые – операторы сочинения внутри сегмента.

3.1.1. Бифункциональность запятой.

Бифункциональность1.

Функции Запятой1 и Запятой2 могут совмещаться, т.е. запятая одновременно оказывается и левой и правой границей. Например, запятая, разделяющая следующие непосредственно друг за другом простые-главные: *Животный страх стучит на машинке, животный страх ведет китайскую правку на листах клочетной бумаги ...* (Мандельштам, далее М) или вторая и третья запятые в цепочке соподчиненных придаточных: *Он*

увидел, что он давно в Москве, что облако подсвечено золотом, что грузовик его стоит ...

Бифункциональность2.

Функции запятой2 и запятой3 могут совмещаться у запятой в ситуации разрыва (третий случай вложения): в S1 в сегмент-хозяин *Процесс одевания, шитье и облачение в шинель ...*, где вторая запятая – запятая3, вложено придаточное, *которому предается Акакий Акакиевич*, относящееся к группе операторов сочинения ограничивать в линейной структуре предложения зоны, где могут находиться слуги сочиненных слов, должно быть вставлено между его хозяином и оператором сочинения. И так бы и случилось, если бы оператором был союз **и**. Т.е., если в сочиненную группу **Процесс одевания и шитье* вставить придаточное *которому предается Акакий Акакиевич*, то в **Процесс одевания, которому предается Акакий Акакиевич, и шитье ...*, вторая запятая осталась бы запятой2, так как есть отдельный оператор сочинения **и**. Если бы строгость записи в естественном языке была такой, как скобочная запись в алгебраических выражениях, то возникла бы запись с двумя подряд стоящими операторами – запятой2 и запятой3: *Процесс одевания, которому предается Акакий Акакиевич,, шитье ...* Однако две запятые ставить не принято, их функции исполняет одна запятая, соответственно возникает ее бифункциональность.

В рассмотренном случае сочетания [**и**] запятая и сочинительный союз **и** оказываются отдельными независимыми знаками с собственными значениями. Но при определении функции запятой необходимо иметь в виду, что запятая в таком и подобных сочетаниях может терять морфологическую автономность. Это происходит, в частности, при сочинении с итерацией сочинительных союзов *и\ или\ да\ либо\ ни\ то\ не\ то\ то ли\ ли:* *Он слышал громкий разговор, и шум ветра, и стучащие по крыше капли дождя...*

Если же нет ни сочинения с итерацией, ни сочинения в разорванном сегменте, то [**и**] чаще всего – бифункциональный знак со значениями запятой1 и запятой2: *Стояло лето Керенского, и заседало лимонадное правительство* (М). Запятая в комбинациях [**а**] и [**но**], может выступать в значениях, близких запятой1\2\3 с аналогичной бифункциональностью.

При этом может возникать стяжение запятой2 и [**и**]\[**а**]\[**но**] в значении запятой1: *Каменноостровский – это легкомысленный красавец, накрахмаливший свои две единственные рубашки, и ветер с моря свистит в его трамвайной голове* (М)

3.2. Тире.

Тире1 и тире2 – аналогично запятой 1 и запятой2 – левые и правые границы любых сегментов: *Он вынул из шкафа визитку и в полном вечернем туалете – безупречно от сандалий до тубетейки*

– с черными шевитовыми ластами на белых ляжках вышел.... (М).

При вложении с разрывом (аналогично левой и правой скобкам) тире - левая и правая границы вводного оборота: *Концертный морозец пробежал по его сухой коже и – матушка, пожалей своего сына – забирался под воротник.* (М)

При этом при вложении без разрыва функцию **тире2**, как и в случае с уточняющим оборотом, берет на себя идущая следом левая граница примыкающего справа сегмента (запятая, точка с запятой, двоеточие) или правые границы предложения (точка, восклицательный или вопросительный знаки), которые при этом оказываются бифункциональны.

Тире3 и Тире4 – левая и правая границы уточняющего оборота, простого или связанного с хозяином отношением гипонимии: *Отчаянно тихо я собрал свои охотничьи принадлежности – сачок, зеленую жестянку на ремне, конвертики и коробочки для поимок – и через окно класной выбрался наружу.* (Н)

Если вложение не порождает разрыва, левая граница следующего сегмента, приняв на себя функцию тире4, становится бифункциональной, как, в частности, точка с запятой в следующем примере: *Я придумывал разные героические способы спасти ее от ее родителей – господина с нафабранными усами и дамы с овальным...лицом; моя мать спросила про них....*(Н)

Тире5 – знак связи подлежащего и сказуемого, т.е. связи слов внутри сегмента, манифестирует проективность сегмента: *Каменноостровский – это легкомысленный красавец, накрахмаливший свои две единственные рубашки, и ветер с моря свистит в его трамвайной голове.*(М). *Ваша первая и единственная любовь – продажная женищина, чистота ее – миф.* (Н)

Тире6 – знак эллипсиса некоторой компоненты при повторе конструкции в одном или разных сегментах. Манифестирует проективность: *Из пяти их сестер Наталья была за Петерсоном, Вера – за Пыхачевым, Нина – за бароном Раушем фон Траубенберг, Елизавета – за князем....* (Н); *Двадцать лет такого танца составляют эпоху, сорок – историю.* (М)

3.3. Двоеточие.

Двоеточие1 – левая граница уточняющего оборота, простого или связанного с хозяином отношением гипонимии. Используется, если обобщающая именная группа стоит слева от уточняющей: *Иван ... беззвучно опустил на нее шелковую клетку: розовый абажур.* (Н) *Из гортани рвутся шипящие, kloкочущие звуки: воинственные «ц», «и», «г».* (М) *Для начала она выкинула всю театральную мишуру: и жар свечей, и красные грядки кресел, и атласные гнезда лож.* (М)

Двоеточие2 – всегда бифункциональная – одновременно правая и левая – граница сегментов:

Очень скоро путем слежки я выяснил, где мать ее снимала дачку: ее скрывала рощица яблоней. (Н)

3.4. Скобки.

Скобки всегда появляются парами, и они всегда – левая и правая границы.

Скобки1 – границы вводного оборота любой структуры, не входящего в граф графов сегментов.

Скобки2 – границы обычных сегментов: *На мокрой простыне, в световом кругу появлялась (судорожно набежав и остановившись) цветная картина...*(Н). *...любуется его смелостью Вася (у которого, по слухам, был дома настоящий револьвер с перламутровой ручкой)* (Н).

3.5. Точка.

Точка1 – конец предложения – правая граница.

Точка2 – знак сокращения.

Неоднозначность членения Т на S возникает, когда контекст позволяет проинтерпретировать точку и первым и вторым способом.

4. Омонимия ЗП и сегментная омонимия линейной структуры предложения.

Хорошо известны факты морфологической омонимии частей речи [8,4,5] и синтаксической омонимии, возникающей на уровне анализа внутрисегментных подчинительных связей в предложении [1].

Сегментная омонимия конкретной линейной структуры S соответствует неразрешимой омонимии ЗП в этом S и может быть выявлена только на уровне снятия омонимии ЗП.

4.1. Примеры неоднозначности сегментации при определении значения запятой – правой границы сегмента vs. оператора сочинения.

Пример1 *...воспоминание, от которого он вздрогнул, когда отец встал, взял чашку и налил в нее чая, открыл форточку и посмотрел на часы.* может быть сегментирован несколькими способами, причем интерпретации синтаксически равноправны и без экстралингвистической информации равно истинны. В данном случае это возникает из-за того, что все сказуемые согласованы и, соответственно, есть несколько комбинаторных вариантов построения сочиненных групп:
Вар1. Третья запятая может иметь значение правой границы, при этом *когда отец встал* – сегмент и, соответственно, разорванный им сегмент – *,взял чашку и налил в нее чая, открыл форточку и посмотрел на часы,*... где четвертая и пятая запятые сочиняют.

Вар2. Третья запятая – оператор сочинения, вложенный сегмент – *когда отец встал, взял чашку и налил в нее чая,* четвертая запятая бифункциональна: правая граница и оператор сочинения. Разорванный сегмент – *...он вздрогнул, открыл форточку и посмотрел на часы,*...

Вар 3. Третья запятая – сочиняет, и четвертая – сочиняет. Вложенный сегмент – *когда отец встал, взял чашку и налил в нее чая, открыл форточку и*

посмотрел на часы, и сегмент-хозяин – он вздрогнул.

Такого рода неоднозначности могут возникать и при построении сочинения других компонент, чаще всего – именных и предложных групп.

4.2. Некоторые неоднозначности интерпретации ЗП при определении левой границы вложенных сегментов, вытекающей из омонимии ЗП.

Значение запятой – левой границы, которое для придаточных S и деепричастных оборотов обычно легко определяется (при условии снятия омонимичных значений подчинительных союзов [5] и деепричастий [4]), может в некоторых случаях составлять проблему.

Пример 1. *Он видел соседа, работающего в огороде отца, брата и деда.* Падежная омонимия Род.\Вин. *отца* позволяет проинтерпретировать первую запятую как омоним [левая граница обособленного причастного оборота vs. оператор сочинения]. Возникает неоднозначность сегментации: Вар.1: все S – один сегмент с сочиненной группой *соседа, работающего в огороде отца, брата и деда*. Вар.2: первая запятая – левая граница причастного оборота, и все S состоит из двух сегментов. При этом омонимия второй запятой [правая граница vs. оператор сочинения] тоже не снимается и порождает два варианта сегментации: Вар.2.1: *Он видел соседа, работающего в огороде отца, брата и деда*, где вторая запятая – граница сегмента и вар.2.2: *Он видел соседа, работающего в огороде отца, брата и деда*, где вторая запятая – оператор сочинения.

Пример 2. *Он заметил в зале своего друга Ваню, известного меломана, нескольких девочек и старика.* Контекстные значения первой запятой – оператор сочинения vs. левая граница приложения. Первая запятая может быть понята и как оператор сочинения и как левая граница приложения. Возникают два варианта сегментации: Вар.1: *своего друга Ваню, известного меломана, нескольких девочек и старика* с одной сочиненной группой. Вар.2: *Он заметил в зале своего друга, известного меломана, нескольких девочек и старика*, где первая запятая левая граница приложения, а вторая – бифункциональна: правая граница и оператор сочинения.

5. Синонимия ЗП.

Рассмотренные выше случаи позволяют говорить о пересечениях множеств значений разных ЗП. Все рассмотренные ЗП могут выступать как левые и правые границы. При этом запятая1 и запятая2 соответственно синонимичны тире1 и тире2, а также скобке левой и правой.

У всех ЗП (кроме скобок и точки) есть синонимичные с т.зр проективности сегмента компоненты значений, например, у запятой3, тире5 и тире6, двоеточия1.

Функциональная синонимия ЗП позволяет унифицировать многие процедуры дешифровки линейной структуры Т.

Список литературы.

- 1) Л.Н.Иорданская. Синтаксическая омонимия в РЯ (с точки зрения автоматического анализа и синтеза). НТИ, сер.2 1967, №5.
- 2) Т.Ю.Кобзарева, Д.Г.Лахути, И.М.Ножов. Модель сегментации русского предложения. Диалог'2001. Аксаково 2001. т.2.
- 3) Т.Ю.Кобзарева. Некоторые аспекты анализа сочинения при сегментации русского предложения. КИИ'2002. Труды восьмой национальной конференции по искусственному интеллекту с международным участием. М.Физматлит т.1.
- 4) Т.Ю.Кобзарева, Р.Н.Афанасьев. Универсальный модуль предсинтаксического анализа омонимии частей речи в русском языке на основе словаря диагностических ситуаций. Диалог'2002 Протвино 2002, т.2 с.с.258-268.
- 5) Т.Ю.Кобзарева. Морфализация in vivo. Труды Международной конференции Диалог'2004.
- 6) Т.Ю.Кобзарева. Принципы сегментационного анализа русского предложения. Московский лингвистический журнал. М. 2004. Т. 8 №1.
- 7) Е.В.Падучева. О порядке слов в предложениях с сочинением: сочинительная проективность. НТИ сер.2, 1971, №3.
- 8) А.Пашенко. Об одном подходе к проблеме снятия омонимии при автоматической обработке текста на естественном языке. НТИ. Сер.2. №4. 1967.
- 9) Ю.Ф.Шрейдер. Свойство проективности языка. НТИ, 1964, №8.