

Глава 4

КОММУНИКАТИВНО-СИНТАКСИЧЕСКИЙ АНАЛИЗ

ТЕОРЕТИЧЕСКИЕ ПРЕДПОСЫЛКИ

Коммуникативно-синтаксический анализ — заключительный этап работы лингвистического процессора, на котором происходит преобразование информации, представленной в виде ПСДО отдельных предложений, в метаязыковую структуру данных.

Неверно было бы думать, что данный этап знаменует переход от синтаксиса предложения к синтаксису текста («макросинтаксису» [Богданов 1977] или «гиперсинтаксису» [Откупщикова 1982]). Противопоставление конструктивного и коммуникативного синтаксиса лишь отчасти соотносится с понятиями предложения и текста соответственно: так, вопросы, связанные с разрешением анафоры, когезией и когеренцией, как правило, обсуждаются применительно к тексту; вместе с тем они могут возникнуть и на уровне предложения. С другой стороны, самые известные коммуникативные значения — тема и рема — выделяются в рамках предложения, однако очевидно, что «на коммуникативную перспективу очередного предложения могут оказывать влияние весьма обширные отрезки предыдущего текста» [Богданов 1977: 135].

К тому же понятие «предложение» традиционно определяется графическими признаками и потому является весьма условным. Так, сложные предложения могут без потери смысла быть разбиты на простые, парцеллированная конструкция может находиться внутри предложения либо быть оформлена последовательностью графически отдельных предложений и т. д. В связи с этим классические задачи лингвистики текста, связанные с выявлением локальных связей между соседними предложениями и выделением таких промежуточных единиц, как сложное синтаксическое целое [Поспелов 1948; Солганик 1973] и сверхфразовое единство [Булаховский 1952: 392–393; Гальперин 1981; Москальская 1981], не являются для нас приоритетными.

Специфика коммуникативно-синтаксического анализа обусловлена не его объектом, а содержательной направленностью. Его ко-

нечная цель — выявление информационной структуры текста на ЕЯ. Ограничения на тип текста сведены к минимуму — в качестве материала могут использоваться любые правильно построенные письменные тексты публицистического, научного, официально-делового стиля, а также художественной прозы. Эта ситуация кардинальным образом отличается от проектов автоматической обработки текстов с заранее известной «денотатной структурой» (ср. [Новиков 1983]). Заметим, что и популярные ныне системы извлечения частичной информации (*information extraction*) также ограничиваются поиском в тексте информации, касающейся заранее известных, заданных объектов и событий¹. Для этого сначала формируются списки соответствующих лексических единиц, затем для каждой единицы создается обобщенный образец ее поведения с точки зрения сочетаемости с другими лексемами (своего рода поисковое предписание). Поиск осуществляется на морфологически размеченных (тэгированных) корпусах текстов. Синтаксический компонент в таких системах, как правило, вовсе отсутствует, а семантика ограничивается заранее заданными списками и шаблонами [Леонтьева 2006: 185–189].

Наша постановка задачи принципиально иная — лингвистический процессор не знает, с каким текстом он работает (стиль, жанр) и о чем в этом тексте идет речь. Он должен уметь преобразовывать результат конструктивно-синтаксического анализа (ПСДО) отдельных предложений в новый формат — информационную структуру текста. Очевидно, что здесь мы вступаем в малоизученную область. Наши разработки в этом направлении носят отчетливо выраженный поисковый характер и не претендуют на решение всего спектра проблем, связанных с анализом текста как единого целого.

На данном этапе мы не можем довольствоваться исключительно сведениями об устройстве языковой системы. Ведь текст — это не очередной уровень языка, расположенный над предложением (как первоначально предполагалось в рамках структуралистской парадигмы), а феномен совершенно иной природы [Бенвенист 1974: 139]. Его организация регулируется не только лингвистическими, но и тематическими, стилистическими, функциональными факторами².

¹ Как правило, дело касается автоматической рубрикации текстов СМИ (ср., например, [Добров 2014]).

² Как писал Роман Якобсон, «в правилах сочетаемости лингвистических единиц существует ведущая вверх лестница свободы. В комбинировании различных признаков фонем свобода отдельного говорящего равна нулю: языковой код уже установил все возможности, которые могут быть реализованы в данном

Следовательно, для анализа текста требуется использовать опыт, накопленный в смежных дисциплинах, прежде всего когнитивных и дискурсивных исследованиях.

Дискурсивные исследования способны пролить свет на особенности структурной организации текстов различных типов. Многообразие речевых жанров и разная степень их структурированности препятствуют обстоятельному изучению этого предмета. Вместе с тем литература по анализу дискурса содержит полезные сведения, касающиеся закономерностей строения отдельных жанров и типов текстов, которые, в частности, могут помочь в интерпретации темпоральной структуры описываемых событий (см. ниже).

Что касается когнитивных исследований, они скорее определяют общую теоретическую платформу наших рассуждений, чем предоставляют конкретные сведения по тому или иному вопросу. Именно здесь, на этапе коммуникативно-синтаксического анализа, становятся актуальными центральные проблемы когнитивной науки, а именно: что люди знают о себе и о мире, откуда они знают то, что знают, как организовано это знание и как оно активируется (пускается в ход) [Gibbs 1996: 29, 40]. Эти вопросы в концепциях различных ученых-когнитивистов получают несколько разное преломление, ср. концептуальную метафору Дж. Лакоффа и М. Джонсона, аспекты образности Р. Лангакера, проблему категоризации, ментальные пространства Ж. Фоконье, оппозицию фигуры и фона и другие схематические системы у Л. Талми и др. (обзор данных концепций см., например, в [Скребцова 2011]). Тем не менее все когнитивные теории языка так или иначе сосредоточены на проблеме так называемого ментального конструирования, осуществляемого говорящим (при переходе от мысли к речи) и слушающим (при обратном переходе от речи к мысли). С точки зрения анализа текста особенно важна позиция слушающего; в то же время не следует пренебрегать и фигурой говорящего, который из доступного набора языковых средств по каким-то причинам выбирает именно те, что представлены в тексте. Последний аспект связан с такой важной категорией текста, как точ-

языке. Свобода соположения фонем в словах очень ограничена, она проявляется лишь в маргинальной ситуации словотворчества. В конструировании предложений и слов ограничений, налагаемых на говорящего, уже значительно меньше. Наконец, при создании текста из предложений действие ограничивающих синтаксических правил прекращается, и свобода каждого отдельного говорящего достигает максимума, особенно если он не должен придерживаться множества языковых клише» (цит. по: [Пеше 1999: 312]).

ка зрения, или перспектива (ср. [Успенский 1970; Шмид 2003: 109–144]), которая, в частности, влияет на интерпретацию дейктических выражений и обуславливает оценочные значения в тексте.

На первый взгляд это может показаться странным, но на данном этапе работы нам близок (если отвлечься от понятийного и терминологического аппарата когнитивизма) тот ракурс рассмотрения языка, его средств и механизмов, который принят в книге классика когнитивной лингвистики Ж. Фоконье [Fauconnier 1985]. Автор строит оригинальную теорию ментальных пространств, призванную описать динамику восприятия дискурса человеком. По Фоконье, этот процесс предполагает непрерывное конструирование мысленных конструкторов: создание и обновление ментальных пространств, введение новых элементов, добавление связей и т. д. При этом внимание исследователя сосредоточено на вопросе о том, какие типы информации, касающейся различных аспектов ментального конструирования, передаются теми или иными языковыми средствами. Не претендуя на исчерпывающее описание, автор отмечает достаточно очевидные случаи (в проекции на английский и французский языки). Например, неопределенный артикль несет информацию о создании нового элемента, грамматические значения времени и наклонения глагола сообщают о статусе ментального пространства, степени его выделенности и актуализованности, анафорические выражения отсылают к уже имеющимся элементам и т. д. Такая постановка вопроса имеет непосредственное отношение к нашей задаче, также связанной с переходом от языковых структур к метаязыковым (правда, иного свойства).

Важно подчеркнуть, что, несмотря на апелляции к психологическим факторам, единственной данностью для нас является текст. Опираясь на результаты конструктивно-синтаксического анализа, мы стремимся выделить из текста последовательность пропозиций и отобразить их на метаязыковую структуру. К сожалению, работа осложняется тем, что, помимо нехватки описаний, ориентированных на компьютерную обработку языка (что характерно и для предыдущих этапов), практически отсутствуют соответствующие практические разработки.

Если все же вернуться к литературе по рассматриваемым вопросам, следует отметить, что на данном этапе работы важными оказываются оба аспекта семантических исследований — как традиционный семасиологический (от формы к содержанию), так и ономасиологический (от содержания к форме), причем послед-

ний представлен в лингвистических исследованиях значительно более скупой, чем первый. В частности, весьма полезны описания грамматики русского языка, организованные вокруг тех или иных семантических полей: упомянем в этой связи серию коллективных монографий «Теория функциональной грамматики» под редакцией А. В. Бондарко, а также детальное описание лексико-синтаксической категории места в [Всеволодова, Владимирский 2009].

ОБЩАЯ КОНЦЕПЦИЯ ИНФОРМАЦИОННОЙ СТРУКТУРЫ ТЕКСТА

Построение информационной структуры текста начинается с создания двух объектов, релевантных для нее в целом и соответствующих, во-первых, фигуре повествователя и, во-вторых, самому тексту. Последнее необходимо для обработки разнообразных метакommunikативных элементов в тексте (ср. [Макаров 2003: 197–201]). Что касается фигуры повествователя, она требуется для того, чтобы каждое предложение (точнее, пропозиция) было привязано к источнику как сказанное либо повествователем, либо персонажем (третий вариант: цитата, атрибутируемая, как правило, некоторому внешнему для повествования лицу).

Разработанная нами концепция информационной структуры текста предполагает выделение двух основных типов данных — объектов и признаков. В целом это разграничение соответствует самому общему (хотя и известному своей неоднозначностью) делению всего того, что составляет пищу нашему сознанию, на предметы и признаки.

Класс объектов включает живых существ (*Петя, писатель, дельфин*), неодушевленные физические предметы (*дом, трамвай, игрушка*), вещества (*молоко, песок*), ментально-информационные объекты (*адрес, доклад*), пространственные объекты, включая географические (*озеро, гора, дорога, Москва, Эверест*) и временные объекты (*вчера, 10 минут назад, I тыс. до н.э., 5 ноября 2017 г.*). Пространственные объекты должны дополнительно учитываться при построении пространственной схемы текста, отражающей взаимное расположение упомянутых в тексте объектов, их перемещение, ориентацию и пр.; причем для географических объектов указываются их координаты. Что касается временных объектов, они служат опорными точками при выявлении темпоральной структуры текста. Это со-

вершенно особенный вид объектов, так что многое, о чем говорится ниже, к таким объектам неприменимо.

Признаки могут характеризовать отдельные объекты или связывать их между собой, а также характеризовать другие признаки (*высокий дом, Петя спит, Петя купил книгу у Васи за 100 рублей, сугубо положительный, быстро бежать*). В отличие от объектов, имеющих пространственную привязку, признакам присущи темпоральные характеристики, которые должны быть отражены в специальном формате (подробнее см. здесь ниже).

Признаки бывают одноместные и многоместные. Признак считается одноместным, если он приписан одному объекту информационной структуры (даже если тот сам по себе является множественным), ср. *умный ребенок; умные дети; рогатый скот; Петя спит*. Многоместные признаки связывают между собой два и более объекта (*Петя учит детей математике; необходимые для работы навыки; разница во времени между Москвой и Камчаткой*). И те, и другие могут обозначать как постоянные признаки (*древний город стоит на берегу реки, грузовой лифт вмещает десять человек*), так и переменные (*голодный ребенок спит на полу, молодая листва шелестит на ветру*).

Признак в совокупности с соответствующими объектами образует пропозицию. Таким образом, согласно нашей интерпретации признака, мы видим пропозицию не только в классических случаях типа *Петя читает книгу, Петя спит*, но и в словосочетаниях *умный Петя, интересная книга* (ср. [Шмелев 1976: 71; Гак 1986: 54])³.

Принципиально важным мы считаем раскрытие так называемых свернутых пропозиций. Прежде всего, речь идет о номинализациях — отглагольных существительных со значением действия (*улучшение, рост, помощь, покраска*). В отличие от личных глаголов, номинализации нередко характеризуются неопределенностью в отношении субъекта действия, а также его временных характеристик, что приводит к незаполненным позициям актантов и дополнительным сложностям при проекции на ось времени⁴. Пустые позиции

³ Строго говоря, можно пойти и дальше, рассматривая в качестве пропозиций соответствующие экзистенциальные пресуппозиции: *Петя существует, книга существует*, однако вряд ли это имеет практический смысл. Появление того или иного объекта в тексте непременно влечет за собой его включение в информационную структуру независимо от его экзистенциального статуса в реальном или воображаемом мире.

⁴ Ср. [Серио 1999].

актантов должны восполняться на основе анализа контекста, но говорить о единой процедуре, похоже, здесь не приходится. Что касается временной отнесенности действия, выраженного номинализацией, и его статуса (реальный / возможный / несуществующий), эти параметры обусловлены грамматическими характеристиками и лексическим значением включающего предиката. Ср.: *в прошлом году обозначился (планировался, провалился) рост производства / в текущем году мы наблюдаем (ожидаем) рост производства / в следующем году будет происходить (намечается, предполагается) рост производства*. Более подробно проблема раскрытия свернутых пропозиций рассматривается в специальном разделе ниже.

Признаковые существительные типа *ум, храбрость, белизна, округлость* также образуют пропозицию (ср. *храбрость солдат, белизна снега*). Их употребление без зависимого существительного, по сути означает, что объект имеет универсальную референцию, ср. *Ум хорошо, а два лучше; Красота спасет мир*.

В информационной структуре признаки связаны между собой различными отношениями — временными, причинно-следственными, условными и т. д. Объекты также могут быть связаны между собой непосредственно, а не только через признаки в тексте (например, *комната — окно, сад — яблоня, поезд — пассажиры* и т. д.).

Такова общая концепция информационного представления. Обратимся теперь к процедуре преобразования синтаксических структур в метаязыковые.

ОТ СИНТАКСИЧЕСКИХ СТРУКТУР К ИНФОРМАЦИОННОМУ ПРЕДСТАВЛЕНИЮ ТЕКСТА

На выходе конструктивно-синтаксического анализа текст представлен в виде цепочки ПСДО предложений — от первого предложения текста вплоть до последнего. Процесс преобразования также идет последовательно слева направо: формирование информационной структуры текста начинается с ПСДО первого предложения, а ПСДО последующих предложений ее дополняют и развивают. Рассмотрим, что собой представляют элементы отдельно взятого ПСДО с точки зрения введенного выше разграничения на объекты и признаки.

Группа подлежащего (П) может содержать как предметную, так и признаковую лексику, соответственно, в информационной струк-

туре требуется создавать как объекты (ср. *мальчик, дерево, стол, линия*), так и признаки (ср. *ум, наводнение, говорить*). В самом простом случае, когда группа подлежащего образована единичным существительным типа 'живое существо' или 'предмет', дело ограничивается созданием в информационной структуре соответствующего объекта⁵. Исключением являются имена реляционной семантики (см. ниже).

Несколько сложнее обстоит дело, если группа подлежащего состоит из двух и более предметных существительных. Тогда наряду с введением новых объектов должны быть эксплицированы связи между ними, ср. *дом Маши, ручка двери, подарок для Пети, поезд из Москвы в Варшаву* и т. д. Вообще, все антропонимы в информационной структуре представлены объектом «человек» с приписанными ему признаками, в числе которых могут быть имя, профессия (*сторож, инженер*), возраст (*старик, подросток*), пол (*девушка, мужчина*), физические, интеллектуальные и прочие характеристики (*слепой, блондинка, дурак, гений, хам, истеричка*), социальные роли (*читатель, пассажир*) и пр.⁶ Так, если группа подлежащего представлена словосочетанием *фильм Тарковского* в информационной структуре создаются два объекта — «фильм» и «человек», при этом последнему приписывается признак «имя: Тарковский». Кроме того, эти два объекта необходимо связать между собой определенным отношением (проблема интерпретации смысловых отношений между членами генитивной конструкции рассматривалась в предыдущей главе). Реляционные имена (*мать, сосед, начальник, одноклассник*), совмещающие функцию идентификации объекта с информацией о его отношении к другому объекту (ранее упомянутому или имплицитруемому), предполагают одновременное добавление и новых объектов, и новых связей. Когда второй объект эксплицирован (*сосед Пети, друг Надиного мужа*)⁷, ситуация более простая. Если же группа подлежащего представлена одиночным реляционным существительным, в качестве второго объекта (своеобразной точки отсчета) выбирается текущий говорящий — повествователь или субъект прямой речи (персонаж). Далее, если мы имеем дело с субъек-

⁵ Кстати говоря, именно объекты образуют костяк информационной структуры — в отличие от синтаксической структуры, где ядром является предикат.

⁶ Из примеров видно, что многие антропонимы требуют квалификации одновременно по нескольким признакам.

⁷ Очевидно, что второй объект, в свою очередь, может заключать отсылку к третьему объекту и т. д.

ектом прямой речи, он, как правило, уже введен в информационную структуру, и референция слов реляционной семантики по умолчанию определяется относительно него. В то же время в ряде случаев (например, в вопросительных предложениях) в качестве второго объекта следует рассматривать фигуру адресата, ср. такие типовые реплики, обращенные взрослым к ребенку, как *Мать-то небось волнуется?* Если же реляционное существительное употреблено вне прямой речи, а повествование ведется от первого лица, отсылка осуществляется к повествователю как объекту с особым статусом (см. выше). Для повествования от третьего лица употребление одиночных реляционных существительных вне прямой речи нехарактерно.

Группа подлежащего может включать существительные, которые в зависимости от контекста выражают предметное или признаковое значение (*врач, преступник, оппозиционер* и т. п.). Соответственно, разработаны правила их интерпретации в различных контекстах, ср. *Врач пришел только вечером* vs. *Петя — врач/учится на врача/работает врачом*.

Если группа подлежащего содержит имена сугубо признаковой семантики (*последний фильм Тарковского, просмотр фильма Тарковского*), соответствующим объектам приписываются признаки, что в совокупности образует пропозицию, пусть даже и с частично незаполненными позициями. Так, в предложении *Просмотр фильма Тарковского занял три часа* номинализация *просмотр* разворачивается в пропозицию «просмотреть (кто? что? — фильм Тарковского)» с пустой позицией субъекта действия, которая предположительно будет заполнена на основании анализа соседних предложений.

Если пропозиция образована исключительно лексемами с признаковой семантикой (*Атака встретила ожесточенное сопротивление; Курить — здоровью вредить; Утром возникает желание позавтракать*), незаполненные позиции неизбежны, однако их интерпретация может быть различной. В первом примере выраженная подлежащим номинализация отсылает к конкретному действию, субъект которого должен быть определен по контексту. Во втором примере подразумевается универсальный субъект, что вообще характерно для одиночного инфинитива в роли подлежащего (так же, как и для одиночного существительного признаковой семантики, ср. *Бедность — не порок, Чистота — залог здоровья*). Что касается третьего примера, то он может иметь оба упомянутых прочтения. Таким образом, можно сформулировать четкие правила интерпретации применительно к одиночным инфинитиву и существительно-

му признаковой семантики и тем самым сузить круг неоднозначных случаев.

Перейдем теперь к сказуемому (С). Сказуемое всегда представляет собой признак (или, по крайней мере, включает его, как в предложениях типа *Это была лисица*), что приводит к численному увеличению связей. Связи соединяют один или более объектов с признаком (*Петя болеет, Маша — филолог, Петя пригласил Машу в театр*). В последнем случае связи маркируются, для того чтобы различать реализуемые семантические валентности (субъект, объект, инструмент и т. д.). Если в качестве актантов выступают признаковые существительные (**Пожар** на атомной электростанции привел к гибели людей), сказуемое устанавливает соответствующие связи между признаками (*пожар, привести, гибель*), а те, в свою очередь, связываются с объектами (*пожар — атомная электростанция, гибель — людей*).

Если предложение содержит однородные подлежащие, информационная структура претерпевает различные изменения в зависимости от типа сказуемого. Возможны следующие ситуации:

- 1) одинаковые признаки приписываются каждому из объектов / признаков (*Рододендроны и бугенвиллеи не растут на севере; Пушкин и Лермонтов — гениальные поэты; Наводнение и пожар — это страшно*);
- 2) объекты / признаки соединяются посредством одного признака, выражающего взаимное действие или взаимное отношение (*Петя и Маша поцеловались / Петя с Машей поцеловались; Гений и злодейство несовместимы*);

Заметим, что примеры типа *Петя и Маша пошли в кино*, допускающие двоякую трактовку (совместное действие или самостоятельные поступки двух людей по отдельности), предполагают построение двух альтернативных структур для учета обоих вариантов.

Дополнения (Д) — это актанты глагола. Как и подлежащее, дополнение может иметь разветвленную структуру и включать слова предметной и признаковой семантики, причем первые не могут подчинять себе вторых (*резать хлеб, сложить вещи в сумку отца*)⁸. Признаковые слова могут быть либо термами (*бороться с наводнением*), либо подчинять себе другие признаковые и предметные имена (*противиться желанию спать, велеть принести чай, осужден за покушение на получение взятки*). Помимо этого, дополнение

⁸ Это возможно разве что в фигуральном смысле, ср. *хлеб изгнанья*.

может представлять собой сентенциальный актант (*Начальник распорядился, чтобы секретарша принесла кофе*). Все эти случаи получают разную интерпретацию в терминах объектов и признаков информационной структуры.

Так, предметные существительные в группе дополнения приводят к созданию новых объектов; если их несколько, то синтаксический хозяин (*сумка отца*) связывается с признаком соответствующим отношением (*сложить в сумку*). Наличие в группе дополнения определений (*сложить в новую сумку, прочитать интересную книгу*) означает прикрепление к создаваемым объектам соответствующих признаков и не меняет коренным образом сути построения.

Другое дело, если сказуемое представляет собой «включающий предикат» [Богданов 1977] и в группе дополнения есть слова предикатной семантики. В ряде публикаций [Семантика и синтаксис... 1981; Типология конструкций... 1985] было показано, что предикаты второго и более высоких порядков отнюдь не редкость в русском языке, к ним относятся фазовые, модальные и многие каузативные глаголы, ср. *начать репетировать/репетицию, должен уехать, заставил извиниться, пригласил совместно отпраздновать наши дни рождения*. Есть и такие глаголы, которые способны замещать свои валентности существительными как предметной, так и предикатной семантики — таковы, к примеру, глаголы зрительного восприятия и оценки [Арутюнова 1976: 136–138]. Как бы то ни было, мы всегда руководствуемся тем правилом, что в информационной структуре предикат должен быть развернут в пропозицию, даже если часть его мест не заполнена. В рассматриваемой ситуации дело осложняется необходимостью определения кореферентных связей между актантами включающего и включенного предикатов. Эта нетривиальная задача была подробно рассмотрена в статьях [Козинский 1985; Шаляпина 1985]. Заметим, что предложения с сентенциальными актантами, как правило, не создают проблем такого рода.

Обстоятельственные детерминанты (О) в зависимости от своего типа требуют разных построений в информационной структуре. Так, обстоятельства места, как правило, требуют создания новых объектов (*на борту самолета, около дома, в мире*), но могут быть представлены событийными именами со скрытой предикативностью, которые должны быть развернуты в пропозицию (*на конференции*). Обстоятельства времени (*во вторник, до Нового года, в период между двумя мировыми войнами*) предполагают создание временных объектов в виде точек, интервалов и т. д. на оси времени.

Конкретная временная локализованность, однако, может замещаться пропозицией: так, выражения *к празднику (город украсили флагами), до съезда (осталось пять дней)* содержат скрытую пропозицию, которая (как всякая пропозиция) имеет определенные координаты во времени и пространстве, но эксплицитных указаний на этот счет текст может и не содержать.

Что касается других видов обстоятельственных детерминантов (цели, причины и т.д.), они по своей сути всегда являются пропозициями, даже если представлены предложно-падежной группой, ср. *из-за дождя* — ‘из-за того, что идет/шел дождь’, *для Петиного выздоровления* — ‘для того, чтобы Петя выздоровел’, *несмотря на отсутствие денег* — ‘несмотря на то, что (у некоторого субъекта) отсутствуют/отсутствовали деньги’. Поскольку такие выражения являются скрытыми пропозициями, возникает уже упоминавшаяся проблема их недостаточной информационной эксплицитности. Незаполненные места предиката могут порождать неоднозначность прочтения. Кроме того, в подобных конструкциях в принципе должны употребляться предикатные существительные (как в примерах выше), но на их место нередко встают существительные с предметным значением, ср. *из-за машины, ради собаки, благодаря Николаю Ивановичу* и т.п. Специфика контекста диктует событийное прочтение, иными словами, подобные существительные также должны разворачиваться в пропозицию, однако проблема информационной недостаточности здесь еще более серьезная, чем в случае номинализаций.

НАКОПЛЕНИЕ ИНФОРМАЦИИ: КОНТРОЛЬ ЗА ОБЪЕКТАМИ И ПРИЗНАКАМИ

Выше мы рассмотрели возможные соотношения между синтаксической структурой отдельного предложения, выраженной набором (-ами) ПСДО, и элементами информационной структуры. Излишне говорить, что обработка каждого следующего предложения (точнее, ПСДО) не только усложняет уже описанные процессы преобразования, добавляя все новые объекты и признаки, но и ставит задачи принципиально иного уровня, связанные с автоматическим анализом текста как единого целого. Нам неизвестны какие-либо лингвистические модели, в которых эта идея была бы осуществлена

в полном объеме⁹. Наши разработки в данном направлении также пока носят фрагментарный характер.

В самом широком плане речь идет о выявлении содержательных связей (локальных и глобальных) между предложениями, которые в совокупности позволяют определять смысловую направленность текста, его тему. Эта задача внимательно исследовалась в рамках лингвистики текста. Как известно, надежды ученых на то, что удастся выделить конечный набор типов связей и способов их выражения, не оправдались. Более того, полемически заостряя ситуацию, Т.М. Николаева [1978] продемонстрировала, что формальная связанность текста (когезия) в общем случае не является ни необходимым, ни достаточным условием его содержательной связанности (когеренции).

Однако когда мы имеем дело с осмысленным текстом, категоричность этого утверждения заметно снижается: наличие когезивных маркеров все же несет определенную информацию о когеренции текста, но ее недостаточно для того, чтобы определить все связи между соответствующими объектами и/или признаками. Следует также подчеркнуть, что выявление средств когезии в том или ином тексте и извлечение из них информации становятся самостоятельной задачей; из разнообразных языковых механизмов когезии (см., например, [Halliday, Hasan 1976]), пожалуй, только лексический повтор и коннекторы не представляют большой сложности с точки зрения автоматической обработки. Что же касается вычисления анафорических и катафорических отсылок, восстановления эллипсиса, установления кореферентных связей между именными группами, определения тема-рематических прогрессий, все это отдельные вопросы, требующие фундаментальной проработки.

Совершенно иной подход к проблеме выявления связности текста был предложен специалистом в области искусственного интеллекта Р.Шенком и его коллегами. Стремясь построить психологически достоверную лингвистическую модель, работа которой будет «насколько возможно, близко соответствовать процессам, происходящим в человеческом мозгу» [Шенк 1980: 20], авторы исходили из известного в когнитивной психологии феномена ожидания. Суть

⁹ Так, знаменитая модель «Смысл↔Текст» на деле работает не с целостным текстом, а с отдельными предложениями [Апресян, Цинман 1998]. Модель Р.Шенка, о которой будет сказано ниже, не была доведена до уровня практической реализации.

его заключается в следующем: по ходу восприятия текста (устного или письменного) человек непроизвольно строит ожидания (предсказания), касающиеся вероятного продолжения текста. Предсказания могут оказаться неверными, но в большинстве случаев они оправдываются. В потоке текста эти предсказания эксплицируют неявные смысловые связи между соседними предложениями, относящиеся к области «наивной психологии» — того, как человек понимает окружающий мир, что побуждает его действовать и т. д.

Моделируя механизм предсказаний, авторы вводят понятие концептуального умозаключения, призванное восполнять подобные «пропущенные звенья». Они выделяют характерные типы умозаключений, которые отражают наиболее регулярные частотные связи между соседними предложениями в тексте и при отсутствии эксплицитных маркеров связи способны помочь компьютеру понимать текст как связанное целое. Подход Шенка служит удачным примером интеграции достижений когнитивной психологии и теории искусственного интеллекта; к сожалению, его реализация оказалась фрагментарной, и модель не получила дальнейшего развития. В то же время идея моделирования концептуальных умозаключений представляется весьма плодотворной в контексте экспликации смысловых связей в тексте. Она могла бы найти свое воплощение в энциклопедическом компоненте, разработка которого является на сегодняшний день насущной задачей.

Обратимся теперь к нашей конкретной задаче — созданию информационной структуры по тексту — и рассмотрим проблему смысловой связанности в этом аспекте. Поскольку элементами информационной структуры являются объекты и признаки, речь идет об определенном отслеживании, контроле за происходящим с теми и другими по мере того, как обрабатываются все новые ПСДО следующих друг за другом предложений. Применительно к объектам могут осуществляться следующие операции:

- создание объекта (операция применяется не только при каузации возникновения нового объекта типа *построить дом, сварить суп, выкопать яму*, но и вообще всякий раз, когда в тексте упоминается объект, на тот момент отсутствующий в информационной структуре¹⁰);

¹⁰ Иными словами, данная операция применяется всегда, когда упоминаемый объект не удается наверняка связать кореферентной связью с уже имеющимися

- создание объекта и его связывание с уже созданными объектами (ср. *Маша — ее муж — его машина — автомастерская, в которой ее чинят — ...*);
- маркировка объекта как несуществующего (начиная с определенного момента), ср.: *человек умер, дом снесли, пирог съели*.

Контроль за объектами в информационной структуре осуществляется путем приписывания им референциальных индексов. Каждому новому объекту либо присваивается новый индекс, либо производится выбор из ранее приписанных индексов. Последнее означает, что объект уже существует в информационной структуре под другим именем, и в этом случае ему просто приписывается еще одно, дополнительное, имя. Эта процедура технически проста, но в теоретическом плане вызывает много вопросов, поскольку затрагивает целый спектр проблем, касающихся референциальных отношений в связном тексте, в том числе: вычисление анафорических отсылок, восстановление эллипсиса, определение референциального статуса именных групп, идентификацию кореферентных лексических единиц. Каждая из них является весьма сложной и многоплановой в аспекте автоматического моделирования обработки текста.

Так, чрезвычайно широкая литература (в особенности зарубежная) посвящена вопросу отыскания антецедентов местоимений (и других близких к ним слов-заместителей) в связном тексте; решение этой задачи, как правило, основывается на использовании вероятностно-статистических методик. Проблема определения референциального статуса имен неоднократно затрагивалась в трудах логиков (преимущественно на материале английского языка); применительно к русскому языку она глубоко и всесторонне анализируется Е. В. Падучевой [1985]. Безартиклевые языки в этом плане представляют еще больший простор для неоднозначности, чем языки с артиклями¹¹.

Впрочем, и в тех и в других языках референциальный статус отдельной именной группы в составе предложения может быть неясен даже читателю, не говоря уже о системе искусственного интеллекта,

объектами. В результате количество объектов в информационной структуре может заметно превышать число соответствующих референтов.

¹¹ Установлению референции может помочь учет некоторых дискурсивных маркеров, однако работы в этом направлении только начинаются (ср. [Борисова 2014]).

так что его определение требует привлечения более широкого контекста, ср. *Скорпион похож на кузнечика* (скорпион вообще или данный скорпион? на кузнечика вообще или на данного кузнечика?), *У кошки глаза зеленые* (у данной кошки или у всех кошек?), *Иван подрался с милиционером* (с конкретным милиционером или каким-то представителем данной профессии?)¹². Если поясняющий контекст отсутствует, в информационной структуре просто создаются соответствующие объекты: *скорпион, кузнечик, кошка, Иван, милиционер*.

В то время как определение анафорических отсылок (и отчасти восстановление опущенных элементов в эллиптических предложениях) базируется на анализе относительно близкого контекста, установление кореференции имен требует семантического анализа в более крупном масштабе. Без решения этой задачи невозможно выявить содержание текста. Использование различных признаков для номинации, в частности, широко распространено среди антропонимов. Так, один и тот же человек может быть обозначен как *лесник, старик, хитрец, толстяк, сосед...* Однако этот ряд контекстуально взаимозаменяемых существительных имеет свои ограничения, и продолжить его словами *подросток* и *тракторист* невозможно (ср. [Шмелев 1973: 106])¹³.

Подобные примеры естественным образом наводят на мысль о необходимости описания совместимости / несовместимости имен с точки зрения кореференции. Именно эта идея была положена в основу гипотезы референциального отождествления имен (см. [Рубашкин 1989: 126–143]). Заметим, что даже если бы эта гипотеза подтвердилась, задача определения кореференции в тексте не была бы решена в полном объеме: ведь помимо отдельных слов, совместимость которых можно задать в словаре, в данное отношение способны вступать сколь угодно сложные словосочетания, не поддающиеся исчислению. Однако, как вынужден констатировать сам автор, его гипотеза была опровергнута контрпримерами. Более того, оказалось, что даже лексический повтор не гарантирует референциального тождества, ср.: *Из пункта А в пункт Б выехала машина, а через 5 часов выехала машина из пункта Б в пункт А* [Рубашкин 1989: 140]. От себя добавим, что не только машины, но и сами пункты,

¹² Примеры Е. В. Падучевой.

¹³ В более широкой перспективе данный вопрос выводит нас на проблему вариативной интерпретации действительности [Баранов, Паршин 1986], а также когнитивные исследования категоризации мира и схематизации описываемой ситуации [Lakoff, Johnson 1980: ch. 24; Talmy 1983; Lakoff 1987; Талми 1999].

условно названные А и Б, могут быть разными: не так уж и сложно найти на карте населенные пункты с одинаковым названием.

Что касается признаков, они всегда приписываются каким-то объектам как опорным точкам информационной структуры. Самая простая ситуация связана с добавлением одностепенного признака к ранее созданному или впервые упомянутому объекту (*Маша ленива; Маша спит*). В случае множественных признаков объекты связываются друг с другом, что может вовлекать уже имеющиеся в информационной структуре объекты, либо требовать создания новых (*Маша любит Петю; Маша разговаривает по телефону*).

Важный вопрос состоит в том, что происходит с признаками, которые были ранее приписаны тому или иному объекту. Оказывается, они могут сохраняться (*маленькая девочка смутилась и покраснела; белый хлеб зачерствел; тетрадь в клетку кончилась; Вася сел на скамейку и углубился в чтение газеты; Иван Иванович шел по улице и думал о друге*), но могут также претерпевать изменения (*яркая краска потускнела и осыпалась; шквальный ветер понемногу ослабевал*) вплоть до отмены (*бедняк разбогател; свежий хлеб зачерствел; Петя долго листал газету, потом выключил свет и уснул*). Пока не вполне ясна природа взаимодействия между старыми и новыми признаками. Хотелось бы каким-то образом систематизировать наши представления о том, в каких случаях признаки изменяют приписанную ранее информацию, а в каких просто ее добавляют. Очевидно, что на это влияет характер обоих признаков — и старого, и нового. Этот вопрос требует углубленного изучения, но уже сейчас можно предположить, что в конечном итоге дело упирается в наши знания не только о языке, но и о мире.

Вопрос о контроле за признаками выводит нас на задачу определения временной структурой описываемых в тексте событий, которая рассматривается в следующем разделе.

ПРЕДСТАВЛЕНИЕ ИНФОРМАЦИИ О ПРОСТРАНСТВЕ И ВРЕМЕНИ

Контроль за объектами, в частности, предполагает отслеживание их местонахождения и перемещения в пространстве. Эта задача включает целый комплекс проблем¹⁴, с которыми в полном объеме

¹⁴ Достаточно полное представление о них можно составить по книге [Кандрашина, Литвинцева, Поспелов 1989].

не смог справиться на настоящий момент ни один коллектив разработчиков моделей искусственного интеллекта. Основная причина, по-видимому, заключается в отсутствии единой системы, в рамках которой может осуществляться описание местонахождения объекта средствами естественного языка. Ведь указание на пространственное расположение может определяться по абсолютной (в Санкт-Петербурге) или относительной шкале (неподалеку от дома), причем в последнем случае отсчет может осуществляться как в экзосцентрической, так и в эгоцентрической системе координат (ср. *В ста метрах на запад от железнодорожной станции Репино* vs. *В ста метрах к западу от моего дома*). Из этого следует, что даже если можно было бы загрузить в модель подробные географические сведения в том или ином формате, это не решило бы проблему интерпретации локативных выражений.

Если попытаться рассмотреть разные типы текстов в указанном ракурсе, мы обнаружим, например, что в коротких новостных текстах преобладает абсолютное шкалирование, хотя нередким является и относительное (экзосцентрическое). Для художественных произведений, напротив, характерно использование относительных шкал; лишь начало описываемых событий может быть привязано к известному, закрепленному в сетке координат, месту. Литературное повествование разворачивается преимущественно в эгоцентрической системе координат, организованной вокруг рассказчика или персонажа.

В связи с этим актуальным становится вопрос об интерпретации дейктических выражений и учете такого параметра, который в различных работах получил название перспективы, или точки зрения, или наблюдателя. Несмотря на традиционное внимание к теме дейксиса со стороны как отечественных (Ю. Д. Апресян, Е. В. Падучева и др.), так и зарубежных (Ч. Филмор, Дж. Лайонз и др.) ученых, заметно усилившееся в последние десятилетия в связи со становлением когнитивной лингвистики и интересом к языковой концептуализации мира [Talmy 1983; Langacker 1991; Яковлева 1994; Булыгина, Шмелев 1997; Логический анализ языка... 1997; Levinson 2003], исследования в этой области, как правило, не носят прикладного характера. Компьютерная лингвистика, по крайней мере отечественная, мало продвинулась в данном направлении.

Правильная интерпретация дейктических выражений важна также для успешной автоматической обработки прескриптивных текстов (инструкций, правил, предписаний, руководств, кулинар-

ных рецептов и др.), многие из которых вообще лишены привязки к абсолютной шкале и широко задействуют эгоцентрическую систему координат. Пространство, к которому осуществляется референция, часто является весьма небольшим (по сравнению, скажем, с художественными и публицистическими текстами), ограниченным рамками того или иного помещения (например, метрополитена), устройства или механизма (например, бытовой техники).

Ограниченность пространства характерна и для дескриптивных текстов, хотя масштаб здесь может быть различным (очевидно, что объектом описания может стать объект сколь угодно малый и, наоборот, сколь угодно большой). Но этот объект заранее задан (как правило, отражен в заголовке текста), и за его границы текст обычно не выходит. Тем не менее это может случиться, если автор сопоставляет описываемый объект с каким-то другим, в некотором роде аналогичном, но расположенном в другом месте. Например, в описание какого-нибудь примечательного здания или сооружения могут быть включены отсылки к другим постройкам того же архитектора, инженера и т.п. В дескриптивных текстах часто встречается первоначальная привязка к абсолютной шкале (ср., например, путеводители). В дальнейшем, правда, такие тексты опираются на относительные шкалы — как экзо-, так и эгоцентрические. Все сказанное, однако, не распространяется на дескриптивные фрагменты, включенные в литературный нарратив (описания персонажей, зарисовки природы и др.).

Заметим, что в русском языке, по-видимому, нет таких локативных выражений (от предлогов до сколь угодно сложных оборотов), которые были бы предназначены исключительно для референции к какой-то одной (из пяти возможных) системе отсчета: абсолютной (которая всегда экзоцентрична), относительной или размытой (последние две могут быть как экзо-, так и эгоцентрическими). В текстах на естественном языке данные системы отсчета причудливо переплетаются, серьезно затрудняя попытки пространственного моделирования. Добавим к этому, что заключенная в тексте информация может касаться различных локативных характеристик — положения одного объекта или взаимного расположения двух и более объектов, расстояния между ними, направления движения, ориентации. Таким образом, необходимо учитывать ряд взаимосвязанных параметров. А если речь идет о динамической ситуации, пространственные характеристики необходимо соотносить с временными координатами, автоматическое определение которых осложнено действием целого ряда факторов (см. ниже).

Дополнительные трудности при автоматической обработке выражений с пространственной семантикой возникают из-за широкого использования в естественном языке размытых систем отсчета [Кандрашина, Литвинцева, Поспелов 1989: 198], которые по своей природе могут быть как экзоцентрическими, так и эгоцентрическими, ср. *недалеко от Петербурга* vs. *недалеко от машины*. Неопределенность локализации, заключенная в выражениях типа *вблизи от, слева от, наискосок от* еще более усиливается за счет того, что оценка расстояния зависит от величины соответствующих предметов: одно дело *чашка рядом с тарелкой*, другое дело — *город рядом с границей*. Имеется и более тонкий момент, связанный с прагматическим параметром «фоновые допущения и ожидания» [Langacker 1988: 67]: скажем, близость работы к дому по-разному будет оцениваться жителями Москвы и маленького провинциального городка.

Известная размытость, неточность семантики присуща не только наречиям и предлогам с пространственной семантикой. Возьмем, к примеру, слово *окрестность*: какую территорию можно назвать *окрестностью* чего-либо? Очевидно, что однозначный ответ здесь невозможен. Во-первых, как и в примерах выше, важен размер так называемой точки отсчета (она же — фон, или ориентир): словосочетание *в окрестностях деревни* предполагает круг с радиусом едва ли больше десяти километров (а скорее даже меньше), в то время как выражение *в окрестностях мегаполиса* способно обозначать гораздо большую территорию. Во-вторых, понятно, что проведение границ «окрестности» при помощи циркуля не соответствует узусу: так, если город расположен близко к границе с другой территориальной единицей (штатом, республикой, страной), его *окрестности* едва ли включают часть этой единицы. Наконец, в данном понятии присутствует смысловой компонент, связанный с доступностью (где-то не более двух-трех часов пути), поэтому можно предположить, что границы «окрестности» с развитием транспортных средств постепенно расширяются.

Напоследок упомянем неоднозначность интерпретации сложных локативных оборотов, включающих запятую, ср. *выступал в Москве, в Большом театре* vs. *гулял на Марсовом поле, на Каменном острове*). Соотношение обозначенных мест, разделенных запятой в приведенных примерах различно: в первом случае второе «вложено» в первое, во втором случае они не пересекаются. Подобная неоднозначность может возникать как в позиции актанта при глаголах с валентностью места, так и в обстоятельственных детерминантах,

относящихся к предложению в целом. Правильная интерпретация таких контекстов невозможна без введения в лингвистический процессор энциклопедического компонента, содержащего описание знаний о мире.

* * *

Если объекты характеризуются положением в пространстве, то признаки — временными параметрами. Определение временной структуры текста на естественном языке является чрезвычайно сложной задачей, что обусловлено рядом факторов. Прежде всего, как говорилось выше, приписывание объекту нового признака само по себе ничего не говорит о сохранении или отмене ранее приписанных признаков. Вероятно, можно провести аналогию с совместимыми и несовместимыми номинациями объектов, хотя здесь мы находимся на еще более зыбкой почве.

Но основная, принципиальная проблема, с которой мы здесь сталкиваемся, — это та, что является единой для отображения пространственных и временных отношений. Мы имеем в виду невозможность построения единой модели, что обусловлено сосуществованием в текстах на естественном языке двух систем отсчета (как пространства, так и времени) — абсолютной и относительной [Кандрашина, Литвинцева, Поспелов 1989: 66–75; Попов 2004: 86, 137].

Применительно ко времени абсолютная система упорядочивает события вдоль оси «прошедшее — настоящее — будущее», где настоящее ассоциировано с моментом речи, прошлое привычно располагается слева от него, а будущее — справа. Следует заметить, однако, что в соответствии со спецификой языковой картины мира настоящее должно моделироваться не как точка на шкале, а как некий интервал с размытыми границами (ср.: *Сейчас я обедаю / занят / разбираю документы / еду на работу / живу в Москве / работаю учителем в школе* и т. п.). Описываемые события на абсолютной шкале могут иметь жесткую привязку к календарному времени (*Стрелки часов были переведены 25 марта в 2 часа ночи по московскому времени*) или требовать дейктического прочтения (*Завтра мы поедem за город*). Примечательно, что настоящее (момент речи) обычно в тексте как таковое не фиксируется, а расположение тех или иных событий слева или справа от него определяется по грамматической форме глагола и обстоятельствам времени, в том числе дейктическим выражениям.

Относительная система отсчета описывает время протекания одного действия (состояния и др.) относительно времени другого действия (состояния). Соответственно выделяются три значения: прямое, предшествующее, последующее. Ср.: *Он появился в тот момент, когда я в нетерпении уже набирал номер его телефона* [прямое]; *Перед тем как идти в гости, мы плотно пообедали* [предшествующее]; *Я заснул около полуночи и проспал 10 часов* [последующее]. Приведенные примеры демонстрируют чистые случаи включения одного события в другое или, напротив, их разделенности; между тем довольно часто встречаются и пересечения интервалов, ср.: *Не успел я толком войти и раздеться, как раздался телефонный звонок; Заря уже занималась, когда они выступили в путь.*

Вообще, временные границы описываемых в тексте событий (как и, собственно говоря, самих событий в действительности) нередко размыты, неопределенны. Чтобы в этом убедиться, достаточно, например, попытаться определить начало (или конец) зари, сумерек, поездки за город (с момента начала сборов? надевания верхней одежды? выхода из квартиры?), рабочего дня, если он официально не задан жестким интервалом (с момента входа в здание? в рабочее помещение? начала конкретной трудовой деятельности?) и др.

Отметим также, что временные объекты весьма разнородны в том, что касается их масштаба (ср. *в эпоху неолита, во II тысячелетии до н.э., 25.09.2017, в 23:14 пятого августа 2000 г.*) и природы (в языке, наряду с обозначениями одного конкретного интервала, имеются темпоральные выражения, референтом которых являются регулярно повторяющиеся промежутки, ср. *по утрам, каждой осенью*).

Еще одна сложность заключается в неоднозначности референциальной отсылки некоторых темпоральных выражений. Так, *в понедельник* может относиться как к прошлому, так и к будущему понедельнику на абсолютной шкале времени, причем данную неоднозначность невозможно надежно разрешить опираясь на форму глагола, ср. *В понедельник я иду на концерт* [будущее] vs. *В понедельник я иду на концерт и вижу...* [прошлое]. Дейктический оборот *в этом году* за исключением приблизительной фиксации на шкале времени сам по себе не несет информации о том, относится ли описываемое событие к прошлому, настоящему или будущему.

Все эти факторы — недискретность внешнего мира, определенная размытость и субъективность его отражения средствами естественного языка, а также неоднозначность референции языковых

выражений — ставят под вопрос пригодность формально-логического аппарата для моделирования содержания текста.

Абсолютная и относительная системы отсчета времени, так же как и соответствующие модели пространственной локализации, в текстах весьма часто совмещаются. В том виде, как они охарактеризованы выше, эти системы способны обеспечивать лишь весьма грубое, упрощенное представление о временной структуре текста на русском языке. В них невозможно учесть все разнообразные аспекты протекания различных действий (процессов, состояний и пр.), выражение которых обеспечивается сложным взаимодействием между лексическим значением глагола, его аспектуальными характеристиками, а также семантикой различных обстоятельственных конструкций. Речь идет о таких параметрах, как различные способы действия, предельность / неопределенность, заполненность / незаполненность временного отрезка действием, единичность / повторяемость действия.

К сказанному следует добавить, что текст может содержать более одной версии событий (действительных или воображаемых) — как в прошлом (значение ирреалиса), так и в настоящем и в будущем. Описывающие их предложения могут опираться на абсолютную или относительную систему отсчета времени, а также на их комбинацию.

Как абсолютная, так и относительная временные ориентации имеют в русском языке специализированные лексические и грамматические средства выражения, однако соответствующие списки (см. [Попов 2004: 190–191]), разумеется, не являются исчерпывающими. Попытки моделирования временной структуры текста на русском языке дополнительно осложняются отсутствием так называемого согласования времен, которое в ряде других языков (английском, французском и т. д.) позволяет более надежно соотносить действия, состояния, отношения и др., выраженные личными и неличными формами глагола. В русском языке время осуществления действия может определяться по отношению либо к моменту речи (*Он придет вовремя*) либо к какому-нибудь другому моменту, в частности ко времени другого действия (*Он думал, что придет вовремя*); проблема состоит в том, что в системе форм времени эти различия не отражены. Иными словами, морфологическое время не отражает непосредственно «семантическое время» [Попов 2004: 191].

ОПРЕДЕЛЕНИЕ ВРЕМЕННОЙ СТРУКТУРЫ ТЕКСТА

Кратко охарактеризовав специфику отражения временных отношений в русском языке и трудности, возникающие при попытке осуществлять автоматический контроль за темпоральной структурой текста, постараемся теперь указать на отдельные подсказки, или зацепки, которые способны направлять этот процесс. В идеале, конечно, хотелось бы маркировать каждый признак некоторым индексом, соотнесенным с временной шкалой (или шкалами), однако на практике дело ограничивается пока обработкой отдельных, более или менее однозначных случаев¹⁵.

Начнем с так называемых прескриптивных текстов, или текстов-инструкций [Хорохордина 2013]. Объединяя широкий спектр речевых жанров (закон, указ, распоряжение, руководство, предписание, приказ, рекомендация, инструкция, рецепт, совет, просьба, порядок действий и др.), этот тип текста обладает характерными особенностями. С точки зрения временной организации интерес представляют следующие.

Во-первых, тексты (или их отдельные компоненты), которые непосредственно описывают некоторую процедуру, демонстрируют иконическое соотношение между порядком следования глаголов и последовательностью описываемых действий. При этом глаголы могут употребляться в форме повелительного наклонения, инфинитива или настоящего времени. Приведем типичные примеры:

*Поместите продукты в духовой шкаф, **выберите** нужный режим духового шкафа и **отрегулируйте** температуру приготовления продукта. Нажимая кнопку, **выберите** режим «Окончание приготовления». Соответствующий индикатор начнет мигать. Затем **действуйте** следующим образом...*¹⁶

*Рыбу **очистить** от чешуи, **отрезать** голову и, не разрезая брюшка, **вынуть** внутренности; **удалить** из головы жабры, **промыть** холодной водой и **нарезать** рыбу поперечными кусками*¹⁷.

¹⁵ Отметим, что в литературе подробно рассмотрены временные соотношения, выраженные последовательностями «глагол СВ + союз и + глагол СВ», «глагол НСВ + союз и + глагол НСВ», «глагол СВ + союз и + глагол НСВ» в пределах одного предложения (см. [Маслов 1984; Падучева 1996: 362–364; Падучева 2016]). Нам неизвестны какие бы то ни было попытки выйти за рамки одного предложения.

¹⁶ Газовый духовой шкаф EOG 23400: Руководство по эксплуатации.

¹⁷ Книга о вкусной и здоровой пище. М., 1952. С. 136.

Смешанную с содой муку **рубят** на столе с маслом. Желтки **растирают** добела с сахаром, **смешивают** с рубленой мукой и **замешивают** тесто, которое **выносят** на 10–15 минут на холод. Охлажденное тесто **раскатывают** в круглую лепешку и **укладывают** в форму или сковороду, смазанную маслом. Поверхность лепешки **посыпают** рублеными орехами, смешанными с сахаром (полстакана), и **заливают** взбитыми в крепкую пену белками, смешанными с сахаром (полстакана). Пирог **выпекают** при температуре 230–240¹⁸.

Догнать и перегнать Америку!

Приди и проголосуй!

Независимо от формы глагола, описываемые действия обращены в будущее — относительно и момента речи (порождения текста), и момента его восприятия адресатом. Настоящее здесь не играет существенной роли: люди могут руководствоваться идентичными инструкциями в разные дни, месяцы и годы. Границы будущего неопределенные, поскольку обусловлены внешними факторами (скажем, принятием новых постановлений, отменяющих старые).

Исключение составляют бытовые директивы, прочно укорененные в «здесь и сейчас» и содержащие конкретные указания относительно действий адресата (как правило, в ближайшем будущем). Можно сказать, что бытовые прескриптивные тексты (скорее речевые акты), такие как просьба, совет, предсказывают возможность исполнения соответствующего действия лицом, к которому обращена директива. (В зависимости от формы глагола этот агент может мыслиться как единичный или множественный, включать или исключать субъекта речи.) В этом случае имеет смысл создавать в информационной структуре соответствующие объекты и признаки, правда, со специальным статусом (не фактивным, а вероятным). Это может оказаться полезным для восполнения недостающей информации (эллипсиса) и экспликации содержательной связанности текста.

Во-вторых, многие прескриптивные тексты включают перечень разрешений и ограничений, выражаемых при помощи отрицательных форм глагола, а также разнообразных модальных слов и выражений, таких как *(не) должен, (не) следует, можно, необходимо, (не)*

¹⁸ Даниленко М. П. Как приготовить дома кондитерские и другие изделия из муки, сладкие блюда, варенье, соки и припасы на зиму (пособие для молодых хозяек). М., 1961. С. 104.

допускается, (не) разрешается, (не) имеет право (-а) не наблюдается хронологического упорядочения соответствующих предикатов: все они одновременно действуют применительно к описываемой ситуации начиная с настоящего (момента порождения текста) вплоть до некой неизвестной точки в будущем. Например:

Гибкий шланг можно использовать при соблюдении следующих условий:

- шланг **не должен нагреваться** выше комнатной температуры (выше 30 °С длина шланга **не превышает** 1500 мм;
- в шланге **отсутствуют** участки сужения просвета;
- шланг **не натянут и не перекручен**;
- шланг **не соприкасается** с острыми кромками или углами;
- шланг **можно легко осмотреть** для проверки его состояния¹⁹.

В-третьих, следует отметить практически полное отсутствие глагольных форм прошедшего времени в прескриптивных текстах.

Наконец, нельзя не упомянуть об особенном употреблении форм настоящего времени при описании потенциальной типовой процедуры, локализованной в будущем (по отношению к моменту составления текста) — так называемом «настоящем предписания» [Кожин, Крылова, Одинцов 1982: 106]. В дополнение к приведенной выше выдержке из кулинарного рецепта сошлемся на тексты официально-делового стиля, где это встречается довольно часто. Вот характерный пример: *О причинении гражданину телесных повреждений в результате применения сотрудником полиции физической силы, специальных средств или огнестрельного оружия полиция в возможно короткий срок, но не более 24 часов, уведомляет близких родственников или близких лиц гражданина*²⁰.

Если не учитывать специфику языка официальных документов²¹, данный фрагмент текста будет (неправильно) интерпретирован следующим образом: объект 'полиция' в информационной базе соединяется с предикатом 'уведомлять', причем последний имеет статус 'реальный, текущий', который соответствует центральному значению форм настоящего времени и приписывается им по умолчанию. Заметим, что в других контекстах такая интерпретация будет верна,

¹⁹ Газовая варочная панель AEG: Инструкция по эксплуатации.

²⁰ Федеральный закон от 07.02.2011 № 3-ФЗ «О полиции», статья 19, п. 5.

²¹ Заметим, что в бытовом разговоре, художественном или публицистическом тексте та же мысль, скорее всего, была бы оформлена в виде условного предложения (*Если гражданин получил телесные повреждения... полиция обязана уведомить...*).

ср.: В данный момент полиция звонит родственникам и уведомляет их о... Это наглядный пример того, что выбор правильной интерпретации морфологических форм глагола зависит от стиля и жанра текста. Проблема в том, что описываемый лингвистический процессор не знает, какой текст он обрабатывает. Так возникает непростая задача автоматического определения типа текста, которой посвящен следующий раздел.

* * *

Возвращаясь к феномену иконической упорядоченности, нельзя не вспомнить о знаменитой статье [Labov, Waletzky 1967], авторы которой обратили внимание на то, что последовательность предложений (точнее, личных форм глаголов) в основной части (так называемом осложнении) коротких устных рассказов отражает порядок соответствующих действий. Применительно к русскому языку их наблюдения могут быть использованы при анализе последовательностей простых предложений²² в нарративном тексте, где сказуемые выражены личными глаголами совершенного вида. Рассмотрим в качестве примера следующий рассказ Д. Хармса:

Однажды Антон Бобров сел в автомобиль и поехал в город.

*Автомобиль **наскочил** на ломаные грабли.*

Лопнула шина.

Антон Бобров сел на кочку возле дороги и задумался.

*Вдруг что-то сильно **ударило** Антона Боброва по голове.*

*Антон Бобров упал и **потерял** сознание.*

На основе анализа подобных примеров было сформулировано следующее рабочее правило: если в тексте употребляются исключительно глаголы совершенного вида в положительной форме и одинаковом грамматическом времени, характеризующие действия одного и того же персонажа, а эксплицитные маркеры темпоральных отношений отсутствуют²³, «по умолчанию» вступает в силу прин-

²² Появление сложноподчиненных предложений нередко нарушает иконический принцип отражения событий в тексте (ср. *Я срочно **выехал** в Москву, как только **получил** письмо*). Что касается сложносочиненных предложений, принцип обычно сохраняет свое действие, если части предложения (клаузы) соединены союзом *и*.

²³ В подобных случаях, как замечает А. В. Бондарко, «глагольный вид выполняет не только свойственные ему аспектуальные функции, но и функции темпоральные» [Бондарко 1971б: 45].

цип иконической интерпретации (ср. известный пример *Veni, vidi, vici* [Якобсон 1983: 107]).

Заметим, однако, что такие тексты встречаются весьма редко — не случайно чистый пример оказалось легче найти в литературе абсурда. В обычном художественном произведении иконичность сохраняется разве что на небольших отрезках текста, ср.:

Утро застало Иванова осунувшимся, похудевшим... Он тихонько пробрался в гостиную, схватил газету и, убежав в спальню, развернул свежий газетный лист.

Через пять минут он вбежал в комнату жены и дрожащими губами прошептал... (А. Аверченко. История болезни Иванова).

Примечательно, что даже если фрагмент повествования насыщен действием, в рассказ часто бывают вкраплены глаголы несовершенного действия, как, например, в отрывке ниже:

Я вздрогнул... Мне почудился скрип отворявшейся двери, потом легкий треск переломанного сучка. Я в два прыжка спустился с развалины — и замер на месте. Быстрые, легкие, но осторожные шаги явственно раздавались в саду. Они приближались ко мне. «Вот он... Вот он, наконец!» — промчалось у меня по сердцу. Я судорожно выдернул нож из кармана, судорожно раскрыл его — какие-то красные искры закрутились у меня в глазах, от страха и злости на голове зашевелились волосы... Шаги направлялись прямо на меня — я сгибался, я тянулся им навстречу... Показался человек... боже мой! это был мой отец!

Я тотчас узнал его, хотя он весь закутался в темный плащ и шляпу надвинул на лицо. На цыпочках прошел он мимо. Он не заметил меня, хотя меня ничто не скрывало, но я так скорчился и сжегился, что, кажется, сравнялся с самой землей. Ревнивый, готовый на убийство Отелло внезапно превратился в школьника... Я до того испугался неожиданного появления отца, что даже на первых порах не заметил, откуда он шел и куда исчез. Я только тогда выпрямился и подумал: «Зачем это отец ходит ночью по саду», — когда опять все утихло вокруг. Со страху я уронил нож в траву, но даже искать его не стал: мне было очень стыдно. Я разом отрезвился. Возвращаясь домой, я, однако, подошел к моей скамеечке под кустом бузины и взглянул на окошко Зинаидиной спальни. Небольшие, немного выгнутые стекла окошка тускло синели при слабом свете, падавшем с ночного

неба. Вдруг — цвет их **стал изменяться**... За ними — я это **видел, видел** явственно — осторожно и тихо **спускалась** беловатая штора, **спустилась** до оконницы — и так и **осталась** неподвижной (И. С. Тургенев. Первая любовь).

Приведенный пример убедительно демонстрирует, сколь невелики фрагменты, удовлетворяющие правилу иконической интерпретации. Появление глаголов несовершенного вида может быть вызвано разными причинами: длительностью соответствующего действия, его повторяемостью, стремлением автора к наглядности (ср. последнее предложение). Оно также может маркировать переход от повествования к описанию (ср.: *Небольшие, немного выгнутые стекла окошка тускло **синели** при слабом свете, падавшем с ночного неба*) — подробнее об этом см. п. «Функционально-смысловые типы речи: описание, повествование, рассуждение».

Следует также подчеркнуть, что присутствие тексте двух и более персонажей и попеременное упоминание об их действиях нарушает единство хронологии. В общем случае для каждого персонажа нужно строить свою шкалу времени. Применительно к приведенному выше отрывку действия рассказчика и его отца можно отразить при помощи двух соотношенных между собой временных прямых, но в других случаях могут потребоваться гораздо более тонкий анализ и сложные построения. Некоторое представление об этом можно составить по следующему отрывку:

Он **надел** круглую шляпу *a la Van Dyck*, блузу, **взял** картон под мышку и **отправился**; я **попелся** вслед за ним. Ася **осталась** дома. Гагин, уходя, **попросил** ее позаботиться о том, чтобы суп был не слишком жидок; Ася **обещалась** побывать на кухне. Гагин **добрался** до знакомой уже мне долины, **присел** на камень и **начал срисовывать** старый дуплистый дуб с раскидистыми сучьями. Я **лег** на траву и **достал** книжку; но я двух страниц не прочел, а он только бумагу **измарал**... (И. С. Тургенев. Ася).

В приведенном фрагменте описываются различные (не коллективные) действия трех персонажей, причем порядок их упоминания в тексте применительно к Гагину и Асе не соответствует последовательности действий этих героев, а именно: Гагин *попросил* прежде, чем *отправился*, Ася *осталась* после того, как *обещалась*. Чтобы правильно выстроить порядок описанных здесь событий, лингвистический процессор должен отождествить действия, выраженные личным глаголом *отправился* и деепричастием *уходя*, связать во

времени уход Гагина, его просьбу и обещание Аси, а также поставить ее обещание раньше того момента, когда спутники оставили ее дома. Отсутствие в русском языке так называемого согласования времен затрудняет отслеживание ретроспекций, что может повлечь ошибки в интерпретации. Вот еще один показательный пример:

Я остановился, запыхавшись, на краю горы и, прислонясь к углу дровяника, стал рассматривать живописную окрестность, как вдруг слышу за собой знакомый голос:

— Печорин! давно ли здесь?

Оборачиваюсь: Грушницкий! Мы обнялись. Я познакомился с ним в действующем отряде. Он был ранен пулей в ногу и поехал на воды с неделю прежде меня (М. Ю. Лермонтов. Герой нашего времени).

Обратим внимание на последний абзац. Знания о мире диктуют читателю то очевидное и единственно верное прочтение, что знакомство Печорина с Грушницким, ранение последнего и его отъезд на воды произошли за какое-то время до того, как персонажи вновь встретились и обнялись. Однако применение принципа иконичности даст обратную (и нелепую) картину: персонажи обнимаются, еще не будучи знакомыми.

Иконичность в тексте не соблюдается также тогда, когда порядок перечисления событий регулируется не фактором времени, а какими-то другими соображениями. Возьмем такой (искусственно построенный) пример: *Иван достиг всего, чего хотел. Заработал кучу денег, удачно женился, построил дом на берегу моря, купил роскошную яхту...* Собственно говоря, по этому контексту нельзя сказать ничего определенного относительно временного порядка «достижений» Ивана, так как непонятно, какими соображениями руководствовался автор текста, выстраивая их в линейную цепочку: хронологией, иерархией ценностей (присущей самому автору, Ивану или адресату) или ходом своих воспоминаний. Суммируя вышесказанное, приходится констатировать, что правило иконической интерпретации статистически оправдано на небольших отрезках текста, но даже там не гарантирует стопроцентный успех.

Легко видеть, что оно не работает на более протяженных фрагментах, где повествование с высокой вероятностью нарушается различными отступлениями, ретроспекциями, рассуждениями, описаниями и т. д. Соответственно, к глаголам совершенного вида добавляются глаголы несовершенного вида (в широком диапазоне форм времени и наклонения), причастия, деепричастия; простые предло-

жения чередуются со сложными. Именно так организовано литературное повествование, а также жанры писем и мемуаров. Подтверждением могут служить даже весьма короткие отрывки текстов, ср.:

*По возвращении из деревни (19 декабря), куда он ездил на короткое время, Пушкин **приехал** прямо ко мне и **жил** в том же доме Ренкевича, который, как сказано, на Собачьей площадке **стоит** лицом, а **задом выходит** на Молчановку²⁴.*

*Однажды Ал. Серг-вич **получает** от Пуцина из Москвы письмо, в котором сей последний **извещает** его, что **едет** в Петербург и очень бы **желал увидеться** там с Ал. С-чем. Недолго **думая** пылкий поэт мигом **собрался** и **поскакал** в столицу. Недалеко от Михайловского, при самом почти выезде, **попался** ему на дороге поп, и Пушкин, **будучи суеверен**, **сказал** при сем: «**Не будет добра!**» — и **вернулся** в свой мирный уединенный уголок. Это **было** в 1825 году, и **провидению** **угодно было осенить** своим покровом нашего поэта. Он **был спасен!**²⁵*

Обратимся теперь к отрезкам художественного повествования, построенных на употреблении глаголов несовершенного вида. Примечательно, что их последовательность также способна иконично отражать порядок действий, ср.:

*В девять часов из Парижа **выходит** специальный поезд, отвозящий в Гавр пассажиров «Нормандии». Поезд **идет** без остановок и через три часа **вкатывается** в здание гаврского морского вокзала. Пассажиры **выходят** на закрытый перрон, **подымаются** на верхний этаж вокзала по эскалатору, **проходят** несколько зал, **идут** по закрытым со всех сторон сходням и **оказываются** в большом вестибюле. Здесь они **сажаются** в лифты и **разъезжаются** по своим этажам (И. Ильф, Е. Петров. Одноэтажная Америка).*

В данном примере порядок следования глаголов несовершенного вида отражает этапы типовой процедуры. Однако чаще всего такие глаголы используются для придания большей наглядности передаваемым событиям, а также для различных отступлений (описаний, рассуждений и др.). В приводимом ниже отрывке, например, пере-

²⁴ Письмо С. А. Соболевского М. Н. Лонгинову // Вересаев В. В. Пушкин в жизни: Систематический свод подлинных свидетельств современников. Т. 1. М., 2001. С. 354.

²⁵ Лорер Н. И. Записи моего времени. Воспоминание о прошлом // Вересаев В. В. Пушкин в жизни: Систематический свод подлинных свидетельств современников. Т. 1. М., 2001. С. 308.

ключение с глаголов совершенного вида на глаголы несовершенного вида знаменует начало ретроспективного рассуждения, а обратный переход сигнализирует возврат к повествованию о событиях, ср.:

*Я назвал себя, и мы **разговорились**. Я узнал, что Гагин, путешествуя, так же как и я, для своего удовольствия, неделю тому назад **заехал** в городок Л., да и **застрял** в нем. Правду сказать, я неохотно **знакомился** с русскими за границей. Я их **узнавал** даже издали по их походке, покрою платья, а главное, по выраженью их лица. Самодовольное и презрительное, часто повелительное, оно вдруг **сменялось** выражением осторожности и робости... Человек внезапно **настораживался** весь, глаз беспокойно **бегал**... «Батюшки мои! не соврал ли я, не смеются ли надо мною», — казалось, **говорил** этот уторопленный взгляд... **Проходило** мгновенье — и снова **восстанавливалось** величие физиономии, изредка чередуясь с тупым недоумением. Да, я **избежал** русских, но Гагин мне **понравился** тотчас (И. С. Тургенев. Ася).*

Важно отметить, что наличие последовательности глаголов несовершенного вида в нарративном тексте (или отрывке) в общем случае не позволяет судить о его темпоральной организации. В качестве примера приведем два коротких фрагмента, принадлежащих перу одного и того же автора: в первом описываются одновременные события, во втором — последовательные, ср.:

*Однажды вечером я **был** дома, в своей одинокой комнате, и **занимался** тем, что **лежал** на диване, стараясь делать как можно меньше движений. Я человек прилежный, энергичный, и это занятие нисколько меня не **утомляло** (А. Аверченко. Мужчины).*

*Навстречу мне озабоченно и быстро **шагает** чиновник Хрякин, торопящийся на службу.*

*Увидев меня, он **расплывается** в изумленную улыбку (мы встречаемся с ним каждый день), быстро **сует** мне руку, **бросает** на ходу:*

— Как поживаете, что подельваете?

*И **делает** движение устремиться дальше.*

*Но я **задерживаю** его руку в своей, **делаю** серьезное лицо и говорю... (А. Аверченко. На улице).*

Анализируя приведенные фрагменты, можно сделать предположение о том, что индикатором темпоральной структуры текста могут служить семантические признаки глаголов (точнее, их ССВ), а именно: поскольку в первом отрывке используются стативные глаголы, вероятно, соответствующие состояния симультанны, а на-

сыщенность второго фрагмента глаголами движения предполагает последовательность, сменяемость соответствующих действий. В дальнейшем будет показано, что в действительности все обстоит сложнее и семантические признаки глагола сами по себе еще ничего не говорят о временной структуре текста: одни и те же глагольные ССВ могут использоваться и в повествовании о событиях, и в описаниях. Однако в ряде случаев, где информация о признаках глагола подкрепляется семантикой наречий и союзов, данное предположение работает. Так, в нижеследующем примере глаголы чувственного восприятия и эмоционального состояния *прислушивалась, глядела, не хотела (вступить)* обозначают одновременные события, а словоформы *заговаривал* (начинательный способ действия), *поднималась* (смена позы), *удалялась, появлялась* (движение), *усаживалась* (смена позы) — последовательные. Временной порядок дополнительно подчеркивается союзом *как только*, наречиями *потом, опять*, ср.:

Она не то, чтобы избегала его... напротив, она постоянно садилась в небольшом от него расстоянии, прислушивалась к его речам, глядела на него; но она решительно не хотела вступить с ним в беседу, и как только он заговаривал с нею — тихонько поднималась с места и тихонько удалялась на несколько мгновений. Потом она появлялась опять, и опять усаживалась где-нибудь в уголку... (И. С. Тургенев. Вешние воды).

Еще одно соображение нуждается в том, чтобы на него обратить внимание. Для правильной интерпретации временного плана нарративного текста, помимо личных и неличных форм глагола, необходимо учитывать свернутые пропозиции: их также требуется локализовать во времени. Показателен следующий отрывок:

Болезнь остановила на время образ жизни, избранный мною. Я занемог гнилою горячкою. Лейтон за меня не отвечал. Семья моя была в отчаянии; но через шесть недель я выздоровел (А. С. Пушкин. Воспоминания).

Если не принимать во внимание предикатное существительное *болезнь*, получится, что определенный образ жизни героя *остановился* до того, как он *занемог*. На самом деле смысловые отношения между предикатами здесь не сводятся к отражению последовательности событий: предложение *Я занемог гнилою горячкою* служит для уточнения характера *болезни*. В рамках нашего подхода подобные случаи

автоматически отслеживаются и адекватно интерпретируются благодаря тому, что лексические единицы соотношены с семантическими признаками вне зависимости от их частеречной принадлежности. В данном случае существительному *болезнь* в семантико-синтаксическом словаре приписано значение 'действие', и на этапе коммуникативно-синтаксического анализа оно трактуется как свернутая пропозиция.

В литературе неоднократно отмечалась сложность и многоаспектность темпоральной структуры художественного текста. Художественное (оно же — сюжетное, событийное) время отличается от реального времени по ряду параметров, а именно: оно может застыть, оно обратимо и хронологическая определенность для него несущественна [Тураева 2009: 23–29]. Поэтому привычно используемая для отражения природы реального времени бесконечная числовая ось (с присущими ей одномерностью, непрерывностью, односторонненностью) не является адекватным инструментом моделирования сюжетного времени. Само только присутствие в тексте более одного персонажа, как правило, предполагает несколько временных планов и, следовательно, требует введения соответствующего числа временных шкал.

Длительность художественного времени определяется длительностью событий фабулы. Сюжет нередко нарушает их хронологическую последовательность. Особенно наглядно это можно видеть на примере детективных произведений, где указанная особенность составляет неперемный атрибут жанра. Темпоральная организация детективного рассказа, а тем более романа отличается высокой сложностью. Фабула является как бы двусоставной, включая линию преступления и линию расследования; в основной части повествования эти линии представлены фрагментарно и непоследовательно, и лишь в конце произведения, как правило, сыщик (в присутствии заинтересованных лиц) осуществляет полную реконструкцию случившегося, т. е. фабула разворачивается в хронологическом порядке (подробнее см. [Скребцова 2012]).

Филологический анализ выявляет присутствие в художественных произведениях несколько типов времени: авторского, событийного, субъективного времени рассказчика и моделируемого условного времени читателя [Николина 2004: 169–170]. Скрупулезно анализируя соотношения между различными видовременными формами русских глаголов и создаваемыми темпоральными планами текста, Н. А. Николина получает в итоге громоздкую систему,

которую, выражаясь языком логики, можно охарактеризовать как много-многозначное соответствие. Излишне говорить, что этот тип связей представляет наибольшие сложности для моделирования в системах искусственного интеллекта.

Руководствуясь суждением М. М. Бахтина о том, что «жанр и жанровая разновидность определяется... хронотопом, причем в литературе ведущим началом в хронотопе является время» [Бахтин 1975: 235], Н. А. Николина пишет, что каждый жанр отличается своим особым соотношением видовременных форм, доминирующей ролью отдельных форм времени. Автор иллюстрирует этот тезис примерами, утверждая, что для афоризмов и сентенций характерно употребление форм настоящего абстрактного или настоящего постоянного, для дневников — взаимодействие глагольных форм в прошедшем перфектном с формами настоящего актуального, для поэтической элегии — сопоставление форм прошедшего и настоящего или будущего, а также финал, подчеркивающий «точечность» сюжета, исчерпанность ситуации или ее неожиданное разрешение, для жанра автобиографии — сопоставление двух временных планов, открытость финала, субъективная сегментация времени и пр. [Николина 2004: 230–262, 267]. Мы намеренно позволили себе это достаточно длинное изложение результатов исследования, чтобы наглядно показать их непригодность, бесполезность для решения задачи автоматической обработки текста. Наше замечание ни в коей мере не оспаривает ценности цитируемой монографии; оно лишь призвано акцентировать тот факт, что имеющиеся описания в подавляющем большинстве имеют филологическую направленность и мало приспособлены к нуждам компьютерной лингвистики.

В нашей практической работе мы сознательно следовали определенным ограничениям. Так, мы не учитывали причастия и деепричастия с точки зрения их локализации на темпоральной оси по отношению к личным формам глагола. Мы ограничились анализом монологических фрагментов, поскольку прямая речь персонажей ставит на повестку дня вопросы, связанные не только с фактуальным содержанием текста (диктумом), но и с его оценкой (модусом), а это выводит нас за рамки текущей темы. Мы также не касались вопроса о сфере действия известного схоластического принципа *post hoc ergo propter hoc*, а именно: в каких случаях при моделировании информационной структуры текста можно ограничиться собственно темпоральными отношениями между событиями, а в каких

требуется добавлять к ним причинно-следственную связь?²⁶ Здесь, как и во многих других затронутых выше аспектах, дело упирается в разработку энциклопедического компонента, описывающего внешний мир, его восприятие человеком и взаимодействие людей с миром и между собой.

Как уже отмечалось, серьезным препятствием на пути автоматической интерпретации темпоральной организации текста является высокий индекс грамматической неоднозначности глагольных форм в русском языке, наличие не только центральных (прямых) употреблений временных форм глаголов, но и многочисленных периферийных (переносных). Так, глагол в настоящем времени может описывать события в прошлом («настоящее историческое») или обозначать будущие действия — намеченные или воображаемые. Формы прошедшего времени могут использоваться для обозначения событий в будущем или в контексте абстрактного настоящего. Формы будущего времени в переносных употреблениях могут отсылать к настоящему (неактуальному) или прошлому²⁷.

Из пяти видовременных форм разве что глаголы несовершенного вида в прошедшем времени практически всегда используются «по назначению», что, впрочем, все равно оставляет простор для некоторой неопределенности (например, по параметру единичности/повторяемости действия, ср. *приходил, спрашивал* и т. п.). Остальные четыре формы «приписаны» к тому или иному временному плану в соответствии со своей прямой функцией, что не мешает каждой из них отсылать к двум другим фрагментам оси времени при переносных употреблениях. Следует заметить, что различные литературные жанры как раз и характеризуются широким диапазоном прямых и переносных употреблений временных форм глаголов. Поэтому с точки зрения темпоральной организации художественное повествование является, по-видимому, наиболее сложным для автоматической обработки типом текста.

²⁶ Ответ на этот вопрос, в свою очередь, обуславливает выбор в пользу событийной или пропозитивной (фактивной) интерпретации. Ср. предложение *С приходом к власти военной хунты в стране начался упадок культуры*, которое допускает как событийное прочтение ('После того как к власти пришла военная хунта...'), так и фактивное ('Вследствие того, что к власти пришла военная хунта') [Арутюнова 1988: 185].

²⁷ Иллюстрации см. в [Бондарко 1971а; Русская грамматика 1980, т. 1: 628–636; Падучева 2010].

* * *

Рассмотрим теперь, как в этом плане устроены другие стили речи — научный и публицистический.

В научных текстах могут встречаться все пять видовременных форм глагола, обозначающих события, которые относятся к прошлому, настоящему или будущему. Как правило, эти формы употребляются здесь в своих прямых функциях. Например, жанр отчета о прошедшем научном событии характеризуется преимущественным употреблением форм прошедшего времени и исходной привязкой к абсолютной шкале отсчета. Вот характерный пример²⁸:

*24 января 2013 г. в Институте русского языка им. В. В. Виноградова РАН (Москва) **прошли** очередные (сорок четвертые) Виноградовские чтения. В этом году они **были посвящены** проблемам изучения истории языка. Организаторы чтений **приняли** решение открыть заседание докладом, посвященным вопросам этимологии.*

*В докладе Ж. Ж. Варбот (Москва) «О соотношении структурно-семантического аспектов анализа в праславянской этимологии» **рассматривается** «вечный» для этимологической методике (поставленный еще в конце XIX в.) вопрос о соотношении (связи / приоритете) одного из двух основных аспектов этимологического анализа: структурного и семантического. Докладчик **предлагает** различать обращение к этим двум аспектам на уровне оценки этимологического решения и на уровне стимула к разработке нового решения. <...>*

*В докладе А. М. Молдована (Москва) «К текстологии 16 Слов Григория Богослова с толкованиями Никиты Ираклийского» **было показано**, что большая часть текстологических особенностей списков древнейшей редакции Слов появилась позднее, когда уже были переведены Толкования <...>*

*С. М. Михеев (Москва) в докладе «О древнерусских глаголических надписях» **рассказал** о выявленных в настоящее время глаголических надписях-граффити на территории Древней Руси <...>*

*С. М. Толстая (Москва) в сообщении «Из истории понятий: “честь” и “честность”» **отмечает**, что концепт “честь” имеет две стороны — “внешнюю” (почтение, уважение и далее — почет, почести, знаки уважения, награда) и “внутреннюю” (высокие моральные качества,*

²⁸ Капитанова Ю. С. Виноградовские чтения 2013 г. // Вопросы языкознания. 2013. № 5. С. 156–157.

моральный кодекс). В русском языке оба понятия **выражаются** словом “честь” и его дериватами <...>

В докладе А. А. Пичхадзе «О взаиморасположении глагольной формы и прямого дополнения в придаточных предложениях в древнерусском языке» **шла речь** о порядке в группе “личная форма глагола + прямое объектное дополнение” в независимых и — подробно — зависимых предикациях. <...>

В докладе П. В. Петрухина (Москва) **рассматривались** древнерусские двойные глаголы — бессоюзные сочетания двух глаголов, обладающие рядом особых семантических, синтаксических и просодических свойств <...>

М. А. Малыгина (Москва) в сообщении «Особенности восприятия греческого текста славянскими переводчиками» **рассмотрела** некоторые ошибки перевода с греческого языка в гимнографических текстах <...>.

Напротив, анонс (о грядущем научном событии), сообщая о том, что состоится в будущем, тем не менее допускает употребление форм настоящего времени у глаголов определенных семантических групп (ср. *планируется, принимаются* и т. д.), ср.:

8–11 мая 2011 г. в Шанхае (Китай) **состоится** XII Конгресс МАПРЯЛ «Русский язык и литература во времени и пространстве». На конгрессе **предполагается** работа по следующим направлениям: государственная языковая политика; роль и место русского языка в государстве и межгосударственном общении; «Русский мир»: русская культура, русский менталитет в условиях современной глобализации; новое в системно-структурном описании современного русского языка; русский язык: диахрония и динамика языковых процессов и др. Заявки на участие (см. ниже) и тезисы объемом до 1 страницы **принимаются** до 31 мая по адресу <...>²⁹

В обоих этих стилевых жанрах правило иконической интерпретации не действует, даже если имеет место последовательность грамматически однородных глагольных форм. В отличие от жанра коротких устных рассказов [Labov, Waletzky 1967], порядок следования предложений определяется здесь не временным фактором, а достаточно жестким шаблоном. Это лишний пример того, что успешная интерпретация текста невозможна без учета типа текста.

²⁹ Respectus Philologicus. 2010. № 17. С. 249.

Хорошо структурированным является также жанр аннотации к научной статье или монографии. С точки зрения темпоральной организации, аннотации обычно представляют собой набор утверждений в настоящем времени, касающихся цели, материала, значимости и основных результатов исследования. Временные рамки соответствующих пропозиций «слева» (на оси времени) ограничены моментом порождения текста, а «справа» не имеют границ. Как и в рассмотренных жанрах отчета о прошедшей конференции и ее анонса, все предикаты здесь являются симультанными: обработка каждого последующего ПСДО приводит к добавлению новых объектов и признаков, не изменяя и не отменяя ранее созданные.

Что касается жанра научной статьи, он сам по себе неоднороден. Так, Е. С. Троянская [1989: 24] выделяет, в зависимости от обсуждаемого предмета и целей авторов, следующие разновидности: 1) статья — краткое сообщение о результатах научно-исследовательских и опытно-конструкторских работ; 2) собственно научная (научно-техническая) статья; 3) передовая статья; 4) историко-научная статья; 5) дискуссионная / полемическая статья; 6) научно-публицистическая статья; 7) рекламная статья.

Т. В. Радзиевская [1986] проводит иное деление, противопоставляя теоретические научные тексты, в основе которых лежит рассуждение, эмпирическим текстам, сводящимся к описанию. Можно утверждать, что последний тип текстов отличается большей однородностью: научные статьи (или их фрагменты), целью которых является описание объекта исследования, строятся преимущественно в настоящем времени. Это же характерно и для соответствующего типа научных монографий, а также недидактических фрагментов учебных пособий. Здесь также все признаки являются симультанными, так что построение информационной структуры данных по тексту сводится в основном к добавлению новых объектов и связей, ср.:

*Если психолога **интересуют** реакции индивида на общественные структуры, а социолог **пытается объяснить** те же самые факты в терминах общества, то для антрополога основной интерес — это ценности социокультурной группы и культурные правила поведения, которые **раскрывают** эти ценности. Иначе говоря, для антрополога каждый акт языкового выбора **представляется** значимым как средство выражения — и постижения — ценностей данной культуры.*

*Антропологический подход, кроме того, **отличается** методикой исследования. Основная цель — изучить нормальное, неконтролируе-*

*мое поведение; основной антропологический метод — включенное наблюдение (participant observation). Антрополог **включается** в жизнь общины, чтобы иметь возможность наблюдать жизнь и поведение людей в нормальных условиях*³⁰.

Прошедшее время в научном тексте используется при обращении к истории предмета или его изучения; при этом возможное чередование с формами настоящего времени не создает композиционного разрыва. Иллюстрируя этот тезис, Д. Н. Шмелев [1977: 139] приводит фрагмент из книги А. А. Шахматова «Очерк современного русского литературного языка». Мы ограничимся отдельной фразой из него, ср.: «Другой процесс *лежит* в основании образования романских языков: вульгарная латынь *усваивалась* народными массами (преимущественно кельтами) и *вытесняла* их, исконно им свойственный язык». Это существенная особенность научного стиля речи по сравнению с художественным [Шмелев 1977].

В теоретических работах, направленных на защиту определенных положений и критику чужих точек зрения, как настоящее, так и прошедшее время могут употребляться при изложении позиций сторон. К примеру, допустимы оба варианта: В «*Логико-философском трактате*» Людвиг Витгенштейн *утверждал/утверждает...* Теоретическая научная статья по своей стилистике сближается с публицистическими текстами, для которых актуальной является задача автоматического разделения фактов и мнений (подробнее см. ниже).

Специфическими чертами научной статьи являются выраженная метаязыковая функция (присущая также учебным текстам) и, как следствие, присутствие так называемого метатекста [Вежбицка 1978] или дейксиса дискурса [Макаров 2003: 198], регулярно реализующегося в конструкциях типа *как мы подчеркивали выше, как будет показано ниже, из сказанного следует, в дальнейшем мы покажем, рассмотрим теперь, см. далее* и т. п. Вследствие этого текст научной статьи оказывается неоднородным: утверждения, направленные на изложение содержания исследования (цель, предмет, процедура и пр.), перемежаются высказываниями, регулирующими способ изложения. Темпоральная организация повествования получается как бы двухслойной: на «обычное» (внешнее) время, связанное с изложением содержания исследования и выражением мнений,

³⁰ Вахтин Н. Б., Головки Е. В. Социоллингвистика и социология языка: учеб. пособие. СПб., 2004. С. 105.

«наслаивается» внутреннее время повествования³¹ (ср.: *начнем с..., покажем теперь, как следует из примеров выше, в заключение подчеркнем...*). Соответственно, возникает задача дифференциации этих двух типов высказываний, каждый из которых может быть представлен широким диапазоном глагольных форм.

Достаточно надежными маркерами использования внутреннего времени служат лексические единицы *выше, ниже, здесь, далее* и некоторые другие в составе сложных слов или в сочетании с глаголами речи, аргументации, демонстрации. Очевидно, однако, что их недостаточно, чтобы выделить все высказывания, связанные с регуляцией способа изложения. Другая группа слов, которые могут служить данной цели, — это формы совершенного вида будущего времени глаголов указанных лексико-семантических групп (типа *рассмотрим, обратимся, перейдем, заметим* и т. п.). Помимо этого, имеет смысл отслеживать употребление прилагательных типа *предыдущий, последующий, дальнейший*, которые, как правило, в научных публикациях употребляются дейктически, хотя случаются и отклонения от этого правила.

Что касается упоминания в тексте научной статьи об авторах (*я, мы, мой, наш, автор*), оно может встречаться как в контексте внутреннего, так и внешнего времени, ср.: *Выше я (мы) указывал на то, что...* (внутреннее время) vs. *В своих предыдущих публикациях я (мы) неоднократно подчеркивал(и), что...* (внешнее время). Более того, некоторые контексты могут оказаться неоднозначными в рассматриваемом отношении, ср.: *Я уже указывал на то, что...* Снятие такой неоднозначности может быть достигнуто при помощи учета графематической информации³²: если предложение сопровождается сноской на научную работу, автор которой совпадает с автором текущей статьи, мы имеем дело с внешним временем, а если сноска отсутствует — с внутренним. Разумеется, разработка подобного алгоритма требует дополнительных усилий, но в принципе осуществима.

* * *

Публицистический стиль охватывает большое число разнообразных речевых жанров. С точки зрения временной организа-

³¹ Оно же — «время адресата», «время читателя» [Николина 2004: 167; 170].

³² Напомним, что результаты обработки текста по мере движения от одного этапа к другому (графематика, морфология, семантика, синтаксис) не теряются: соответствующая информация накапливается и в любой момент может быть использована во всем объеме.

ции газетные тексты могут оказаться не менее сложными, чем художественные. Если в художественной литературе проблемы, возникающие при автоматической обработке, в значительной степени обусловлены частыми переходами от повествования к описанию и обратно, в текстах СМИ существенным осложняющим фактором является присутствие аргументативного начала. Практически все публицистические жанры (за исключением разве что хроникальной и короткой информационных заметок), наряду с рассказом о событиях, подразумевают более или менее эксплицитное выражение различных мнений, принадлежащих как третьим лицам (названным или подразумеваемым), так и самому автору. Пропозиции в тексте могут относиться как к событийному содержанию текста, так и к его оценочному компоненту.

Здесь мы сталкиваемся со сложной проблемой автоматического разграничения фактов и мнений. Информационная структура должна включать оба эти компонента, но учитывать принципиальное различие между соответствующими типами данных. В описываемом лингвистическом процессоре акцент сделан на событийном компоненте, а оценки прикрепляются к соответствующим пропозициям. Формат представления мнений в настоящее время находится в стадии разработки. Предполагается некоторым образом их систематизировать посредством введения шкалы или закрытого списка значений, что, конечно, приведет к известному огрублению, но зато позволит эффективно обрабатывать большие массивы текстов. Помимо этого, оценочный компонент должен включать ссылку на автора и тем или иным способом локализоваться на шкале времени.

В связи с недостаточной разработанностью данной проблемы обратимся к рассмотрению того жанра публицистического стиля, который лишен (эксплицитной) оценочности, а именно короткой информационной заметки. Несколько упрощая положение вещей, будем считать, что в ней вовсе нет оценочного компонента, а содержание исчерпывается изложением конкретных событий.

Отличительными особенностями данного жанра являются сильно выраженная привязка к моменту создания и опора на абсолютную систему отсчета времени. В дискурсивных исследованиях отмечается, что новостные тексты СМИ структурированы по принципу релевантности, т. е. наиболее важная информация представлена в начале сообщения [Ван Дейк 1989: 132–133]. В связи с постоянным обновлением содержания газетных новостей их темпоральная организация демонстрирует что-то наподобие обратного хронологиче-

ского порядка повествования: сначала идет самая свежая информация, отражающая события на момент создания репортажа, а затем предыстория (фон), разворачивающаяся назад, ко все более ранним стадиям развития сюжета. В литературе это иногда называется «структурой перевернутой пирамиды» (см., например, [Bell 1998]).

Заметим, однако, что приведенные суждения относятся к основной части информационной заметки. Что касается рамки, в которую она заключена, первое предложение, как и заголовок (если он есть), обычно содержит формулировку темы всего сообщения, в то время как последнее (или последние) возвращает читателя в настоящее (момент создания сообщения). Получается, что в целом этот жанр похож на бытовой нарратив, который изучал У.Лабов (см. выше в данном параграфе), за тем исключением, что в основной части можно наблюдать действие «обратного иконического принципа». Покажем это на следующем примере:

***На Васильевском острове угнали машину
Управления делами президента России***

На Васильевском острове угнали машину Управления делами президента России.

*Как передает АН «Оперативное прикрытие», 2 мая в 5:20 утра в районную полицию **обратился** заместитель директора ФГУП «Кремлевский». По словам 43-летнего чиновника, около 1:30 он **обнаружил** пропажу служебного автомобиля BMW-740 с московскими номерами. Иномарка **была припаркована** у дома 30 по улице Кораблестроителей.*

Размер ущерба устанавливается. Машина оборудована штатной сигнализацией, которая была включена³³.

Однако не следует возводить структуру «перевернутой пирамиды» в ранг универсального правила. Она действительно работает, в случае если новостной текст повествует о заметном происшествии, его возможных причинах и последствиях. Вместе с тем жанр короткой информационной заметки, выделяемый по формальному критерию «10–30 строк» [Лысакова 2005: 115, 119], в содержательном плане является неоднородным. Другой вид информации подразумевает иные логико-смысловые отношения между предложениями и иную темпоральную структуру текста. Проиллюстрируем наше суждение заметкой, включающей выражение мнения и его обоснование, ср.:

³³ URL: <http://www.gazeta.spb.ru/1634263-0/> (дата обращения: 26.03.2018).

Эксперт: Российские товары не смогут заменить импортные

Прилавки магазинов в результате запрета на импорт товаров из Евросоюза скорее опустеют, чем «обрусуют», считает ритейлер, управляющий партнер Management Development Group Дмитрий Потапенко. Об этом он заявил в интервью в студии телеканала РБК.

«Я с большим восторгом слушаю отечественных производителей, когда их спрашивают журналисты или люди с улицы: у вас отечественный товар? Он говорит: да, отечественный. Я говорю: хорошо, давайте табличку разложим. Оборудование чье? Импортное — немецкое, голландское. Семенной материал чей? Голландский, канадский, и далее по списку. Для того, чтобы говорить “отечественный товар”, нужно сначала лет двадцать возвращать ученых, которые займутся семенным материалом, вложить в них кучу денег, они будут долго ошибаться, и только после этого у нас появится что-то свое, — заявил эксперт. — Что мешало последние двадцать лет поддерживать нашего сельхозпроизводителя?»³⁴.

Нетрудно видеть, что в данном тексте отсутствует структурирование по принципу «перевернутой пирамиды». Другая разновидность газетного сообщения, в котором также не наблюдается иконической упорядоченности (ни прямой, ни обратной), представлена текстами, где общий тезис, вынесенный в заголовок, и первое предложение далее конкретизируются частными утверждениями о фактах, подкрепляемыми цифрами, ср.:

В некоторых регионах РФ из-за санкций дорожают продукты

После введения Россией эмбарго на поставки продуктов из ЕС и США в стране начали расти оптовые цены на некоторые товары в Москве и регионах. Об этом пишет газета «Коммерсантъ».

Так, по данным издания, в российской столице свинина для мясоперерабатывающих комбинатов подорожала на 6%. В Приморье некоторые виды рыбы подорожали на 40%, а в Сахалинской области рост цен на куриные окорочки достиг 60%.

Об этом на заседании совета по продовольственной безопасности правительства Сахалинской области заявил региональный министр сельского хозяйства, промышленности и торговли Николай Борисов.

³⁴ URL: <http://www.rosbalt.ru/video/2014/08/07/1301485.html> (дата обращения: 26.03.2018).

По его словам, в Южно-Сахалинске и близлежащих городах оптовые цены, к примеру, на сыры выросли на 10%, а на мясо — на 15%.

В Приморье мясо подорожало на 26%, импортируемые из Китая яблоки — почти на треть. На острове розничные цены на продовольствие увеличились на 3–60%, поясняют представители торговых сетей Сахалина³⁵.

Итак, несколько неожиданно, жанр короткой информационной заметки оказался внутренне неоднородным с точки зрения типов содержательных связей между заголовком и основным текстом, а также способов развертывания темы. Отсюда следует невозможность постулирования общей для данного жанра временной структуры текста.

* * *

В рамках данного раздела мы стремились показать разнообразие темпоральной организации в текстах различных жанров и неоднородность видовременных форм глаголов в русском языке. Цель этого аспекта анализа предложения заключается в том, чтобы вскрыть и описать те формальные признаки, которые можно использовать при построении алгоритма автоматического определения временной структуры текста. Речевые жанры широко варьируют по степени схематичности, и только очень немногие укладываются в определенный ясный шаблон³⁶. Представляется, что рассмотренный материал (примеры литературного, научного и публицистического повествования, прескриптивных и дескриптивных текстов) убедительно демонстрирует сложность задачи.

АВТОМАТИЧЕСКОЕ РАСПОЗНАВАНИЕ ТИПА ТЕКСТА

Понятие типа текста

В предыдущем параграфе мы сделали попытку охарактеризовать хотя бы в грубом приближении особенности временной организации текстов отдельных жанров и соотнести ее с интерпретацией грамматических форм глагола. Следует признать, что в наших рассуждениях было некоторое лукавство: ведь тем самым предполагалось, что жанр обрабатываемого текста заранее известен. Вместе с тем в пре-

³⁵ URL: <http://www.24rus.ru/more.php?UID=113839> (дата обращения: 26.03.2018).

³⁶ К шаблонам относятся прежде всего жанры официально-делового стиля, оформление которых регулируется нормативными документами.

дисловии к настоящей монографии мы подчеркивали, что лингвистический процессор не знает, с каким текстом (стиль, подстиль, жанр) он работает. Поэтому было бы некорректно обойти стороной вопрос о том, каким образом возможно осуществлять автоматическое распознавание типа текста. Как и в случае с определением темпоральной структуры, речь идет о различных видах информации, приписанной на более ранних этапах анализа, которая может служить своего рода «опознавательными знаками» того или иного типа текста.

Прежде чем обсуждать эту тему, следует пояснить содержание используемого здесь понятия «тип текста». Как будет видно из дальнейшего изложения, в зависимости от ситуации оно может соответствовать функциональному стилю, подстилю, отдельному жанру или даже функционально-смысловому типу речи в рамках жанра (ср. чередование повествования и описания в художественном произведении, описания и рассуждения в научной или публицистической статье и т. д.). Предвосхищая упреки, связанные с широтой и размытостью введенного понятия, мы апеллируем к соображениям практического удобства, важным в контексте прикладного исследования. Выражение «тип текста» позволяет избежать обсуждения широкого спектра проблем, связанных с классификацией текстов по различным основаниям (общее представление об этом можно составить, например, по работам [Beaugrande 1980: 197–199; Богданов 1993: 9–17; Филиппов 2003: 183–205], номенклатурой функциональных стилей, подстилей и жанров русской речи, соотношением понятий стиля и жанра (см., например, [Орлова 1997]), набором конституирующих признаков речевых жанров, их обусловленностью коммуникативной ситуацией и пр. Эти сложные и спорные вопросы, относящиеся к теории функциональной стилистики, на протяжении многих десятилетий дискутируются в соответствующей литературе (см., например, [Жанры речи 1997; 1999; 2002]), однако однозначного решения до сих пор не получили, и едва ли оно возможно в принципе. Наша задача, напротив, вполне конкретная, и заключается она, как уже было сказано, в том, чтобы постараться обнаружить достаточно надежные дифференциальные признаки однородных в стилистическом отношении текстов, с тем чтобы обеспечить их успешную автоматическую обработку³⁷.

³⁷ Ср.: «Понятие “тип текста” в настоящее время принято как рабочий термин в современных исследованиях по теории текста, в частности в лингвистике текста. Обозначает он эмпирически существующие формы манифестации текстов.

Сама идея поиска отличительных признаков того или иного стиля (жанра) не является новой для языкознания, в том числе отечественного. Начиная с 1960-х годов в нашей стране проводились соответствующие исследования, правда, многие из них имели выраженную филологическую или учебно-методическую ориентацию, так что результаты не представляют большой ценности для компьютерной обработки языка. В это же время, однако, наблюдался рост интереса к применению количественных методов в языкознании и складывались различные отрасли лингвистической статистики, в том числе стилеметрия, основной (и исторически первичной) задачей которой является атрибуция текстов. Последняя, как правило, предполагает установление авторской принадлежности произведения, когда оно анонимно, подписано псевдонимом или в случае литературных мистификаций (ср. [Марусенко 1990]). Вместе с тем существует и более широкое понимание атрибуции, ср.: «Наряду с авторской выделяют не авторскую атрибуцию, цель которой — отнесение данного текста к определенному языку, стилю, периоду времени, литературной школе, литературному направлению, идейному течению, общественному классу и т. д.» [Мартыненко 1988: 6].

В области квантитативной типологии текстов, становление и расцвет которой связан с периодом 1960–1980 гг., были получены некоторые результаты, касающиеся формальных критериев дифференциации различных типов текстов, таких как средняя длина слова, предложения, относительная частота словоупотреблений по частям речи и пр. [Статистичні параметри... 1967; Статистика речи... 1973; Алексеев 1983; Лингвистические распределения... 1985]. Следует заметить, что в рамках данной задачи (в отличие от предшествующих) мы допускаем использование количественных методов. Это связано с тем принципиальным обстоятельством, что текст не является очередным (следующим за синтаксическим) уровнем языковой структуры³⁸. Его построение регулируется не только правилами

Расхождения в толковании понятия “тип текста” еще достаточно велики. Оно трактуется то слишком узко, то слишком широко...» [Валгина 2003: 112].

³⁸ Ср. знаменитые слова Э. Бенвениста: «...нужно признать, что категорематический уровень включает только одну форму языковых единиц — предложение. Оно не составляет класса различимых единиц, а поэтому не может входить составной частью в единицу более высокого уровня. Предложение может только предшествовать какому-нибудь другому предложению или следовать за ним, находясь с ними в отношении последовательности. Группа предложений не образует единицы высшего уровня по отношению к предложению. Языкового уровня,

языка, но и внешними, экстралингвистическими факторами³⁹, следовательно, собственно лингвистического анализа может оказаться недостаточно.

Рассуждая об отличительных признаках стиля, Д. Н. Шмелев [1977: 80] пишет, что «различные функционально-стилистические разновидности совсем не в одинаковой степени характеризуются именно особыми, присущими только им языковыми средствами. Одни из выделяемых разновидностей характеризуются использованием по существу всех средств языка, другие в большей или меньшей степени ограничены в этом отношении». Самым замкнутым функциональным стилем является официально-деловой стиль, на втором месте — научный стиль. Публицистика, по словам автора, отличается нестабильным репертуаром «собственной» лексики и фразеологии и почти неограниченной возможностью привлечения самых разнообразных языковых средств. Применительно к языку художественной литературы едва ли можно говорить о наличии каких-либо специфических языковых средств [Шмелев 1977: 80-81].

Из сказанного следует, что и возможности распознавания стилей различны: чем он более однороден, тем легче его узнать и наоборот. Сложнее всего обстоит дело с идентификацией художественного стиля; скорее всего, ее следует проводить «по остаточному принципу». Но говоря в целом, следует признать, что различия, как правило, не столь существенны, чтобы на них можно было уверенно опираться при определении типа текста. Дело осложняется и тем, что речевые жанры, их структурные и языковые особенности исследованы фрагментарно, несистемно и явно недостаточно для того, чтобы можно было воспользоваться готовыми описаниями. Требуется огромная предварительная работа по анализу разнообразных типов текстов и выработке соответствующего процедурного представления.

Обратимся теперь к краткому обзору имеющихся в литературе наблюдений, касающихся различных типов текстов в интересующем нас аспекте. Руководствуясь суждением М. М. Бахтина [1986] о том, что каждый жанр характеризуется особой тематикой, композицией и стилистикой, постараемся обнаружить соответствующие призна-

расположенного выше категорематического уровня, не существует» [Бенвенист 1974: 139].

³⁹ Это обстоятельство усиленно подчеркивали видные представители дискурсивного анализа (см., например, [Coulthard 1977: 4, 7; Stubbs 1983: 102–103]), стремясь преодолеть заблуждения, распространенные на ранней стадии развития лингвистики текста.

ки для разного рода текстов, при этом делая основной упор на выявление именно языковых (лексических, морфологических, синтаксических) различий.

Научный стиль

Наиболее обстоятельно в литературе описан научный стиль речи, чьи структурные и языковые особенности получили отражение в ряде монографий и коллективных сборников (см., например, [Язык научной литературы 1975; Особенности... 1976; Язык и стиль... 1983; Кожина 1966; 1972; 1986; Лариохина 1979; Лингвистические особенности... 1981; Ванников 1984; Научная литература... 1985; Разновидности... 1989]), не говоря уже о множестве статей. Имеются работы, посвященные конкретным жанрам научной речи — научному реферату [Паташюте 1977; Закономерности... 1982; Грязнухина 1985; Радзиевская 1986], рецензии [Синдеева 1982; Кудасова 1983; Ляпунова, Мешков, Терехова 1990; Красильникова 1999], обзору [Воробьева 1986], статье [Клейменкова 1989; Вагурска 1992], аннотации [Яцко 1993]⁴⁰. Хотя большинство работ датируется еще доперестроечным периодом, основные характеристики жанров сохраняются и в наши дни, что объясняется высокой консервативностью научного стиля (по сравнению, скажем, с публицистическим).

Среди общих черт научного стиля исследователи отмечают почти исключительное использование формы 3-го лица глагола, заметное преобладание форм настоящего времени, сравнительно высокую долю существительных среднего рода, явный перевес родительного падежа (и прежде всего родительного приименного) над прочими падежными формами; в синтаксическом плане — широкое использование пассивных и безличных конструкций, предложений с составным именным сказуемым, сложных предложений союзного типа [Кожин, Крылова, Одинцов 1982: 94–99].

Широкомасштабное исследование [Кожина 1966] было направлено на то, чтобы при помощи скрупулезных статистических подсчетов выявить специфику научного текста по сравнению с художественным (из всех стилей наиболее далеко от него отстоящим). Однако полученные М. Н. Кожинной количественные различия в ос-

⁴⁰ Приводимые перечни, разумеется, не претендуют на полноту, в частности не упоминается обширная учебно-методическая литература, посвященная выработке навыков создания первичных и вторичных документов научного стиля речи на родном и иностранном языках.

новном не превышают нескольких процентов, а потому едва ли могут служить достаточным основанием для разграничения этих стилей (ср. [Комарова 2004: 161]). Причина тому — высокая неоднородность научного стиля речи, включающего большое число разнообразных жанров: письменных и устных⁴¹, монологических и диалогических, первичных и вторичных, обладающих открытой или, напротив, жестко фиксированной структурой.

Помимо собственно научных жанров, образующих ядро данного стиля (монография, статья, диссертация, доклад и пр.), выделяются также справочно-энциклопедические, научно-оценочные, научно-учебные, научно-деловые и другие типы текстов, относящиеся к периферийным и пограничным областям [Троянская 1989]. Более того, внутри отдельного жанра (например, научной статьи) могут проводиться дальнейшие подразделения и выделяться частные разновидности [Троянская 1989: 24]. Наиболее ярко черты научного стиля выражены в ядерных жанрах; остальные несут на себе отдельные характеристики смежных стилей. Так, в жанре научной рецензии, близком к публицистическому стилю, сильно выражено личностное начало и широко используется оценочная лексика [Красильникова 1999]. Научно-деловые жанры (описание изобретения, патент, авторское свидетельство) несут на себе черты официально-делового стиля, что проявляется прежде всего в их синтаксических особенностях.

С учетом вышесказанного можно предположить, что анализ однородного в жанровом отношении материала способен дать весьма убедительные результаты. Действительно, в коллективной монографии [Закономерности...1982], посвященной статистическому анализу текстов научных рефератов, в частности, содержатся следующие впечатляющие утверждения: «...из 20 возможных в системе спряжения глагольных синтетических форм — инфинитив, шесть личных форм настоящего времени, шесть личных форм будущего времени, четыре формы прошедшего времени, три формы повелительного наклонения — употребляются (с относительной частотой не менее 0,40 % каждая) лишь восемь, т. е. 40 %. Их суммарная частота составляет 99,9 % всех употреблений глагольных словоформ. При этом только три глагольные формы — инфинитив, 3-е лицо единственного и множественного числа настоящего времени — охватывают около 97 % всех употреблений глагола, в том числе фор-

⁴¹ Напомним, что устная речь не рассматривается в рамках данной работы.

мы настоящего времени 77,23 %...». Ценность подобной статистики в контексте автоматического распознавания типа текста бесспорна. Правда, далеко не все жанры характеризуются наличием столь устойчивых инвариантных признаков; к тому же данное исследование является заметным исключением в широком ряду работ по функциональной стилистике.

Если даже опора на грамматическую информацию не позволяет уверенно распознавать научную речь, тем более это справедливо в отношении лексики. Достаточно надежным показателем, по-видимому, может служить только высокое содержание терминологических единиц, относящихся к соответствующей дисциплине, ср.: «Ведущим признаком этого (научного — *Авт.*) стиля нужно признать терминологичность словарного состава» [Гальперин 1965: 69]. В то же время это явление может наблюдаться и в публицистике, ср. аналитические статьи по экономике и финансам в деловых изданиях. Что же касается так называемой общенаучной лексики (*метод, критерий, структура, анализировать* и т. п.), ее присутствие в тексте тем более не гарантирует его принадлежности к рассматриваемому стилю.

Вообще, как показано в исследовании А. И. Комаровой [2004], неоднородность, «разножанровость» научной речи влечет наличие, наряду с инвариантными, вариативных признаков регистра. Допустимыми становятся даже «стилистически маркированные научные тексты», содержащие идиоматические выражения, метафорические употребления слов, экспрессивные лексические единицы. В целом эти черты не свойственны научной речи, но полностью исключить их появление в тексте нельзя.

Представляет интерес замечание Л. В. Красильниковой [1999: 14] о невозможности употребления в научной речи местоимений *ты* и *вы* и, соответственно, глагольных форм второго лица. Казалось бы, это наблюдение могло бы быть успешно использовано для автоматического отграничения научного стиля от художественного и публицистического. В то же время следует заметить, что оно не распространяется на произведения научно-популярного характера⁴², в которых вполне могут употребляться языковые средства адресации (в особенности если текст предназначен для детской аудитории), ср.: *если вы помните, как вы знаете, вы можете возразить, видите, по-*

⁴² Как известно, статус научно-популярной литературы остается неопределенным: разными исследователями она понимается то как отдельный стиль, то как подстиль. Обсуждение этого вопроса не входит в рамки данного исследования.

нимаете и пр. [Нистратова 1997]. Помимо этого, указанные местоимения и глагольные формы вполне можно встретить в цитируемых источниках, а также языковых иллюстрациях (в лингвистических текстах). Правда, и те и другие имеют графическое выделение (кавычками или шрифтом), но для того чтобы это отследить, требуется учитывать этот вид информации — наряду с лексической и грамматической, которые традиционно попадают в поле внимания исследователей.

В этой связи мы предлагаем оригинальный подход к распознаванию центральных жанров научной речи (монография, статья, диссертация), основанный на учете графематической информации, приписанной атомам, токенам и их группам на начальном этапе анализа. Наша идея заключается в том, что отличительной особенностью этих жанров является большое разнообразие используемых символов, как буквенных (различных алфавитов), так и небуквенных: достаточно вспомнить оформление ссылок на литературу, сноски, аббревиатуры, цитаты, формулы, маркированные и нумерованные списки, библиографию и пр. Вести учет типов символов возможно благодаря отличительной особенности описываемого лингвистического процессора, связанной с накоплением информации (графематической, морфологической, семантической, синтаксической) по ходу анализа, при движении от одного этапа к другому. Тот факт, что на самом последнем и наиболее сложном (коммуникативно-синтаксическом) этапе оказывается востребованной графематическая информация (начальный этап анализа), представляется нам, помимо прочего, свидетельством правильности выбранной идеологии построения лингвистического процессора⁴³.

Официально-деловой стиль

Официально-деловой стиль речи, обслуживающий сферу административно-правовой общественной деятельности, сближается с научным стилем по ряду параметров — сложному синтаксису, рас-

⁴³ Развивая эту тему, можно отметить роль тире для идентификации художественного текста на русском языке. Как известно, тире в начале абзаца используется для обозначения реплик персонажей в художественном тексте (в английском языке аналогичную роль выполняют кавычки, у которых более высокий индекс неоднозначности). Поэтому если русский литературный текст насыщен диалогами, благодаря тире в такой специфической позиции он будет успешно распознан как таковой.

пространенности безличных и пассивных конструкций, широкому использованию терминов. Если в научном тексте оценка и экспрессия заметно ослаблены по сравнению с художественным и публицистическим стилями, в официально-деловой речи они, можно сказать, сведены к нулю. В проекции на нашу задачу — построение информационной базы по тексту — это означает, что снимается задача выявления мнений и их носителей.

В литературе отмечается, что текстам официально-делового стиля свойственно точное следование «норме» изложения: объективно-нейтральной (в языковом плане) и линейно-последовательной (в плане композиции) подаче содержания [Одинцов 1980: 139–141]. Норма развертывания текста для многих письменных жанров данного стиля является настолько устойчивой, что существуют фиксированные схемы изложения определенного содержания и готовые стандартные формы не только языкового выражения отдельной мысли, но и развертывания целого текста [Татарникова 2006: 94]⁴⁴. Это, несомненно, облегчает автоматическое распознавание подобных текстов⁴⁵, а также обеспечивает однозначность их интерпретации (ср. [Мурзин 1991: 20–21]).

Лексические и грамматические черты жанров официально-делового стиля с разной степенью детализации описаны в ряде работ, ср. [Кожин, Крылова, Одинцов 1982: 104–108; Веселовская 1990; Гуртуева 1991; Кыркунова 1996; Арефьева 1997; Анисимова 2000; Татарникова 2006].

Ядром официально-делового стиля является законодательный подстиль, характеризующийся внутренней однородностью. Для языка законов (указов, постановлений и т. п.) характерны нейтральность изложения, отсутствие экспрессивно-эмоциональных языковых средств, широкое использование абстрактной лексики, юридической терминологии, высокий уровень формализации. Безличный характер законодательной речи исключает использование в ней личных местоимений, личных форм глаголов 1-го и 2-го лица, повелительного и сослагательного наклонений, вопросительных и воскли-

⁴⁴ Именно среди документных жанров мы встречаем наиболее жестко построенный тип текстов — то, что Шварцкопф [1996: 272] называет «текстом-матрицей».

⁴⁵ Не следует также забывать, что все официальные документы непременно снабжены заголовком, определяющим их жанр (ср. *договор, справка, заявление, приказ, акт, протокол* и т. д.).

цательных предложений. Практически исключено использование глагольных форм прошедшего времени.

Важной особенностью законодательного текста, облегчающей его автоматическое распознавание (но существенно затрудняющей конструктивно-синтаксический анализ!), является сложная композиционная структура, подразумевающая многоступенчатое разбиение на части, разделы, главы, статьи, пункты. Таким образом, и в этом случае формальная (графическая) информация оказывается полезной для идентификации типа текста.

Публицистический стиль

Публицистические тексты в силу своего многообразия также имеют мало черт, которые бы позволили автоматически отличать их от произведений иных стилей. Трюизмом стало известное высказывание В.Г. Костомарова [1971] о чередовании экспрессии и стандарта в газетном тексте, однако не вполне ясно, как это наблюдение можно использовать в контексте автоматического анализа. В качестве другой отличительной особенности публицистической речи обычно упоминается обязательность авторской оценки. Однако это утверждение нуждается в уточнении. Во-первых, как указывалось выше, в официально-информативном подстиле (жанрах хроникальной и короткой информационной заметки) она практически сведена к нулю. Во-вторых, проблема автоматического выделения мнений в компьютерной лингвистике еще далека от решения, причем это справедливо даже в отношении первичной задачи, связанной с распознаванием компьютером того факта, что текст несет в себе оценку. Не следует также забывать, что субъективная модальность широко присуща также художественным и научным текстам.

Как замечает О.В. Дунаевская, «в рамках публицистического стиля есть формы, тяготеющие к официально-деловой речи (заметка), есть — к научной (статья), а есть — и к художественной (очерк, эссе, фельетон); итак, мы видим: функционально-стилистическая “сетка” накладывается на отдельно взятый функциональный стиль, воспроизводя свои черты словно в уменьшенном варианте» [Дунаевская 2010: 60–61]. Вероятность обнаружения специфических черт отдельных жанров выше, чем стиля в целом, хотя и не столь высока, как обнаружилось применительно к отдельным жанрам научной и официально-деловой речи. Это наглядно видно на примере описания [Коньков 2004], посвященного информационным и аналитическим

жанрам газетной речи. По-видимому, сохраняет справедливость и суждение А. Н. Васильевой [1982: 29] о значительном расхождении между жанровым и лингвостилистическим членением.

Примером исследования отдельных жанров публицистической речи может служить статья [Карпилович 2013], где автор обращается к рассмотрению структурных и языковых особенностей проблемной статьи (а также проблемного очерка и интервью). Т. П. Карпилович выделяет присущую данному жанру суперструктуру (в терминах Ван Дейка [1989]), а также ряд лексических маркеров и характерных синтаксических конструкций. Так, она констатирует, что в надзаголовке, заголовке, подзаголовке или в начальных абзацах текста часто присутствуют слова *проблема, вопрос, препятствие, трудность*; при изложении сути проблемы нередко используются отрицательные предикаты, слова с отрицательным значением, которые описывают различного рода запреты, невыполнение каких-либо действий, приводящие к конфликту. Впрочем, автор отмечает необязательность указанных лексических средств. По ее мнению, большой оптимизм с точки зрения возможностей формализации всеяют средства языкового воплощения другого прототипического компонента проблемных очерков и интервью, а именно «пути (способы) решения проблемы». Его индикаторами в первую очередь служат структуры с прямой или косвенной речью, а также характерные лексические комплексы, включающие слова со значением специальностей и должностей, к примеру: *специалист фирмы, начальник департамента по социальному развитию* и т. п. Именно в этой структурной части статьи широко представлена эмотивная лексика, императивные конструкции, в том числе с отрицательными предикатами, риторические вопросы, наречия со значениями важности, необходимости и прочие средства выражения оценки.

Эти наблюдения представляют интерес, однако возможность их использования для автоматической идентификации произведений рассматриваемого жанра вызывает сомнения (сама автор статьи преследует иную цель, а именно автоматическое реферирование). Прежде всего, они связаны с тем, что, «в отличие от композиции, компоненты суперструктуры часто не присутствуют в тексте в виде отдельных законченных фрагментов, а находятся в различных частях текстового пространства» [Карпилович 2013: 6]. Это фактически сводит к нулю ценность выделения структурных компонентов и характерных для них языковых средств. Таким образом, в «сухом остатке» мы имеем некоторый необязательный перечень типов

лексических единиц и синтаксических конструкций, и трудно сказать, насколько надежно они могут служить индикаторами рассматриваемого жанра. Безусловно, это объясняется спецификой публицистического жанра в целом и никоим образом не может считаться недостатком обсуждаемого исследования.

* * *

Наиболее сложную задачу представляет собой идентификация художественного текста. По-видимому, в отношении него в общем случае⁴⁶ следует применять принцип «от противного», а именно: гипотеза о принадлежности текста к этому стилю выдвигается тогда, когда проверка по всем чертам, характерным для текстов других типов, дает отрицательный результат (т. е. они либо не представлены вовсе, либо соответствующие подсчеты не превышают средних значений).

***Функционально-смысловые типы речи:
описание, повествование, рассуждение***

В зависимости от темы сообщения и коммуникативных намерений речевая деятельность организуется так, чтобы логико-понятийные свойства сообщаемой информации соответствовали намерениям говорящего.

Тема сообщения, являясь существенным и необходимым признаком всякого связного текста, составляет ядро содержания текста и определяет его структуру. Тематику конкретного текста можно определить по объекту высказывания, ответив на вопрос «о чем здесь говорится?».

Коммуникативные намерения говорящего различаются по способам изложения информации, называемыми композиционно-речевыми формами или функционально-смысловыми типами речи [Нечаева 1974, 1999; Брандес 1990; Всеволодова 2000; Валгина 2003: 75–96] (Стилистический энциклопедический словарь... 2006). В тексте их можно определить, ответив на вопросы «как излагается информация?», «о чем говорится?», «какие информационные задачи решает автор?». В зависимости от ответов на эти вопросы выделяется три функционально-смысловых типа: описание, повествование и рассуждение.

⁴⁶ За исключением текстов, насыщенных диалогами (см. выше), а также драматургических произведений.

Каждый из названных типов речи⁴⁷ имеет отличительные черты по речевым свойствам используемого арсенала языковых средств (функционально-семантические признаки) и применяемым логикориторическим способам и приемам изложения информации (понятийно-смысловые признаки).

С логической точки зрения типы речи дифференцируются следующим образом: для сообщения характерны диахронические отношения между признаками объекта (развивающиеся, изменяющиеся действия или состояния), для описания — синхронные отношения (одновременные признаки объекта), для рассуждения — каузальные (причинно-следственные отношения между суждениями об объекте).

В функциональном и смысловом отношении описание сводится к ответам на вопрос «какой объект обсуждается?», «какие стороны объекта обсуждаются?». Тип описание означает, что в тексте содержится фактическая констатирующая характеристика сосуществующих объектов, свойств предметов, наличных явлений и ситуаций. Описание используется для статического изображения действительности, для констатации состояния действительности, местности, интерьера, важных черт личности, характера или внешности человека.

Типологическим подтипом описания является разъяснение. Разъяснение используется говорящим для истолкования содержания какого-либо текста, объяснения совокупного смысла какого-либо события, какой-либо ситуации на основе личного опыта и обусловленного знаниями истолкователя.

Как в описании, так и в разъяснении используются по преимуществу логические и аналитические средства языка, обращенные к разуму (к рассудку). Композиционная схема описания выглядит следующим образом:

- 1) указание на объект описания (тема описания);
- 2) выделяемые денотативные признаки объекта описания;
- 3) характеристика свойств, оценка качеств объекта описания.

⁴⁷ Помимо этих трех способов изложения, известных еще из старинных риторик, современные авторы выделяют и другие, дополнительные — объяснение [Вейзе 1985], доказательство [Цветкова 1982]. Еще больше различных типов речи предлагается в работе [Одинцов 1980]; расширение основного списка поддерживает и Н. С. Валгина, указывая на то, что обозначенных выше трех типов оказывается явно недостаточно, когда мы выходим за рамки художественных произведений [Валгина 2003: 92–93]. Так, в диалоге нередко вопросы и побуждения к действию. Наш подход выводит такие случаи за рамки понятия функционально-смысловых типов речи, и мы считаем возможным ограничиться классическими тремя типами.

Повествование, отвечая на вопросы типа: «что происходит или произошло с объектом повествования?», «какие события повлияли на объект повествования?», отражает действительность в динамике, характеризует объекты, события во временной последовательности и развитии. Коммуникативные намерения, реализующиеся в описаниях и повествовании, обычно облекаются в форму дескриптивных высказываний. Композиционная схема повествования следующая:

- 1) начало события, в которое вовлечен объект информации, завязка;
- 2) развитие события;
- 3) кульминация события;
- 4) конец события, развязка.

В основе рассуждения лежат ответы на вопросы: «почему выбран данный объект рассуждения?», «почему данный объект таков?», «почему нечто произошло/происходит с объектом рассуждения?», «каковы причины изучаемого объекта, явления (события, действия)?», «каковы следствия из данного явления?», «что значит для нас изучаемый объект, изучаемое явление (событие, действие)?».

Этот функционально-смысловой тип речи используется при описании причинно-следственных связей явлений и событий действительности, с его помощью производится обобщение фактов, устанавливаются имплицативные отношения между событиями и их рациональной оценкой. Рассуждения составляют базу нормативных высказываний, т. е. устанавливающих правила и ценности. В рассуждении как функционально-смысловом типе речи используются по преимуществу логические и аналитические средства убеждения, обращенные к разуму (к логосу).

Рассуждение строится по следующей композиционной схеме:

- 1) тезис, т. е. суждение, обосновываемое в процессе аргументации;
- 2) доказательства, аргументы, подтверждаемые примерами;
- 3) вывод или заключение (может отсутствовать в тексте).

Очевидно, что многие речевые жанры допускают чередование разных типов речи⁴⁸. Так, в научных текстах часто встречается опи-

⁴⁸ Считаем нужным подчеркнуть, что здесь и далее выражение «типы речи» следует понимать именно в вышеуказанном смысле, поскольку в литературе встречаются и другие его трактовки, ср. [Шмелев 1977; Кожин, Крылова, Одинцов 1982; Солганик 2010: 80–94].

сание объекта и рассуждение, но может присутствовать и повествование (например, если речь идет о процедуре экспериментального исследования). В публицистике, за редким исключением, обязательным компонентом является рассуждение, но оно, в свою очередь, базируется на некоем описании объекта или сообщении о событии. В художественном произведении чаще всего перемежаются сообщение и описание, но встречается и рассуждение.

Рассмотрим, что представляют собой данные типы речи с точки зрения построения информационной структуры текста. Начнем с описания (дескриптива). Согласно О. А. Нечаевой, описание — это «модель монологического сообщения в виде перечисления одновременных или постоянных признаков предмета в широком понимании» [Нечаева 1974: 94]. В таких текстах референтная функция отчетливо доминирует над остальными [Чебанов, Мартыненко 1999: 6]. В литературе выделяются разные виды описания в зависимости от типа ремы. Наиболее распространенными являются описания с качественной и предметной ремой [Валгина 2003: 80–85].

Описания с качественной ремой статичны: внимание автора фиксируется на свойствах одного предмета, который является темой. Такого рода описания нередко встречаются в научно-популярных изданиях, справочниках, энциклопедиях, ср.:

Марс — четвертая от Солнца планета Солнечной системы. Она хорошо заметна невооруженным глазом как весьма яркая звезда, имеющая максимальную звездную величину $-2,0^m$. В своем видимом движении Марс довольно быстро перемещается по небу, поскольку его орбита расположена близко к Земной.

Диаметр Марса вдвое меньше земного, а его масса в десять раз меньше массы нашей планеты. Тонкая марсианская атмосфера сильно разрежена. Это позволяет наблюдать поверхность планеты с помощью телескопов. На ней хорошо различимы темные и светлые участки, горы, ущелья, а также знаменитые полярные шапки⁴⁹.

Другой характерный пример описания с качественной ремой заимствован нами из путеводителя:

Родос, главный город острова, существует на самой его северной окраине с 408 г. до н. э. С трех сторон город омывается морем. Вдали почти всегда ясно видно побережье Малой Азии. Длинный галечный пляж, часто обдуваемый сильными ветрами, окаймляет западное по-

⁴⁹ Планеты и созвездия: иллюстрированный справочник. Вильнюс, 2013. С. 94.

бережье. На севере почти вплоть до городского центра простирается песчаный пляж⁵⁰.

Описаниям с качественной ремой соответствуют следующие построения в информационной структуре. Создается объект — тема текста (в наших примерах — *Марс, Родос*), и ему последовательно, по мере обработки предложений, приписываются признаки. (Для нас пока достаточно весьма общего, грубого представления, не учитывающего различных нюансов построения, которые могут быть связаны, например, с тем, что разные признаки могут характеризовать объект не целиком, а какие-то его отдельные части и аспекты.)

Описания с предметной ремой также предполагают статичное положение наблюдателя, который как будто смотрит вокруг себя и приписывает характеристики отдельным объектам, составляющим «компоненты» того места, в котором он находится, ср.:

В то лето я жил в маленьком северном городе. Город стоял на берегу реки. По реке плыли белые пароходы, грязно-бурые баржи, длинные плоты, ширококулые карбасы, с запачканными черной смолой бортами. У берега стояла пристань, пахнувшая рогожей, канатом, сырой гнилью и воблой. На пристани этой редко кто сходил, разве только пригородные колхозники в базарный день да унылые командировочные в серых плащах, приезжавшие из области на лесозавод.

Вокруг города по низким пологим холмам раскинулись леса, могучие, нетронутые: лес для сплава рубили в верховьях реки. В лесах попадались большие луговины и глухие озера с огромными старыми соснами по берегам. Сосны все время тихонько шумели. Когда же с Ледовитого океана задувал прохладный влажный ветер, нагоняя тучи, сосны грозно гудели и роняли шишки, которые стучались о землю (Ю. Казаков. Арктур — гончий пес).

Такие описания предполагают иные построения в информационной структуре, чем те, что были указаны выше для описаний с качественной ремой. Здесь теме соответствует некий «суперобъект» — в данном случае *город*. Ему подчинены прочие объекты — *пароходы, пристань, леса, озера, сосны* и пр., каждому из которых приписывается свой (-и) признак (-и)⁵¹. Встает задача моделирования

⁵⁰ Родос: путеводитель (серия «Nelles Verlag»). СПб., 2010. С. 17.

⁵¹ Проблема выделения подобных «суперобъектов» в тексте и построения соответствующих иерархических связей в информационной структуре далека от

пространственных связей между всеми этими объектами. Требуется также включить в эту картину фигуру повествователя, которая в нашей информационной структуре наделена особым статусом.

Независимо от типа ремы все создаваемые признаки имеют статус 'одновременный'. Что касается того, следует ли считать эти признаки постоянными, заслуживает внимания наблюдение В. А. Яцко о различиях между описанием в научной и художественной речи. По его словам, в отличие от описания в научном тексте, описание в художественном произведении характеризуется тем, что «наряду с постоянными, повторяющимися признаками объекта зачастую указываются и временные признаки, непосредственно наблюдаемые в момент речи...» [Яцко 1998: 129–130].

Важная особенность описания как типа речи — нейтральная модальность, невозможность выражения модуса мнения [Яцко 1998: 127]. Следовательно, здесь не встает задача распознавания мнений (*opinion mining*) и привязки пропозиций к носителям соответствующих точек зрения. Правда, сказанное безоговорочно верно лишь для научного описания: в публицистике и художественном тексте сам выбор слов, осуществляющий категоризацию действительности, несет эмоциональную нагрузку, так что здесь можно говорить о другой, более тонкой и сложной проблеме — определении тональности текста (*sentiment analysis*).

Рассматривая языковые индикаторы описания как типа речи, сошлемся на наблюдение О. А. Нечаевой, которая пишет: «Описанию не свойственна косвенная, ирреальная модальность (со значением пожелания, условия) в форме сослагательного наклонения или значение побуждения, просьбы, т. е. повелительное наклонение, наконец, невозможна и предписывающая модальность в виде необходимости, долженствования» [Нечаева 1974: 44]. Весьма редко используются в описании и вводные слова [Нечаева 1974: 45]. В целом «ограниченное использование в описании вводных слов, как и наклонений гла-

решения, и дело не упирается здесь исключительно в разработку энциклопедического компонента, содержащего знания о мире. Дело в том, что многие отношения, которые в тексте оказываются иерархическими, не являются таковыми в действительности: так, *пароходы, плоты, вобла, колхозники* не составляют непрерывный атрибут городского пейзажа. В принципе, никакой энциклопедический компонент не способен вместить в себя все то, что может быть упомянуто в тексте, когда речь идет о городе. Сказанное, разумеется, справедливо и в отношении других объектов: посредством ассоциативных связей практически все может быть связано со всем.

гола, по-видимому, объясняется одной причиной: констатирующей сущностью описательной речи» [Нечаева 1974: 46]. Вместе с тем, как видно из приведенных формулировок, речь идет не о жестких правилах, а закономерностях, и сама автор на страницах своей книги приводит примеры, идущие вразрез с указанными тенденциями.

Из характерных синтаксических особенностей описания как типа речи следует отметить номинативные, безличные и эллиптические предложения, характеризующие место, время, состояние окружающей среды и пр. [Нечаева 1974: 82–90]. Впрочем, эти филологические наблюдения, к сожалению, не подтверждены статистически.

Заслуживает внимания упоминание о том, что художественное описание чаще строится в прошедшем времени, а научное или деловое — в настоящем [Нечаева 1974: 50]. Впрочем, это справедливо скорее в отношении двух последних стилей. В художественном же описании нередко сочетаются формы настоящего и прошедшего, глаголы несовершенного и совершенного вида. Иными словами, такие формальные показатели, как смена глагольного вида или времени, не являются надежным индикатором смены типа речи (см. также ниже). В качестве иллюстрации используем пример из цитируемой монографии [Нечаева 1974: 52]:

*Бричка бежит, а Егорушка **видит** все одно и то же — небо, равнину, холмы... Музыка в траве **притихла**, старички **улетели**, куропаток не видно. Над поблекшей травой... **носятся** грачи: все они похожи друг на друга...* (А. П. Чехов. Степь).

Пространные описания демонстрируют весьма прихотливое переплетение временных пластов и, соответственно, личных форм глагола. Это справедливо и в отношении тех художественных произведений, которые целиком построены в соответствии с данным типом речи. Таков, к примеру, рассказ И. А. Бунина «Антоновские яблоки». Вот один лишь абзац:

*Из такой трепки сад **выходил** почти совсем обнаженным, засыпанным мокрыми листьями и каким-то притихшим, смирившимся. Но зато как красив он **был**, когда снова **наступала** ясная погода, прозрачные и холодные дни начала октября, прощальный праздник осени! Сохранившаяся листва теперь **будет висеть** на деревьях уже до первых заморозков. Черный сад **будет сквозить** на холодном бирюзовом небе и покорно ждать зимы, пригреваясь в солнечном блеске. А поля*

уже резко **чернеют** пашнями и ярко **зеленеют** закустившимися озими... Пора на охоту! (И. А. Бунин. Антоновские яблоки).

* * *

Сообщение (повествование, нарратив) рассказывает о некоторой последовательности действий или состояний, поэтому объектам в информационной структуре должен автоматически приписываться статус 'временный'. Далее встает вопрос об отношении данного признака к предшествующим и последующим признакам того же объекта. Отменяет он предыдущий (*поплакал и успокоился*) или нет (*лег и уснул*)? Отменяется ли он последующим? Строго говоря, поскольку анализ осуществляется поступательно, достаточно всякий раз задаваться лишь первым вопросом, т.е. «оглядываться назад», соотносить данный признак с предшествующим, уже имеющимся в базе (если он есть).

Сам по себе это вопрос непростой, но нужно понимать, что основная проблема лежит гораздо глубже. Ситуация, в принципе, напоминает ту, с которой мы столкнулись ранее, рассматривая темпоральную организацию различных речевых жанров (см. параграф «Определение временной структуры текста»). Мы исходили из того, что жанр нам известен, и лишь позднее признались, что распознавание жанра — самостоятельная сложная задача. Так и здесь: мы начали говорить о временных границах признаков в нарративе, но откуда лингвистический процессор знает, что имеет дело с нарративом? Единственное, что доступно наблюдению, — это языковые формы, от них и следует отталкиваться. Перестроим наши рассуждения в этом ключе.

Фундаментальная проблема заключается в том, что нарратив в русском языке может задействовать разнообразные грамматические формы глаголов (как совершенного, так и несовершенного вида), причастий, деепричастий, а также существительные признаковой семантики — отсюда трудность распознавания этого типа речи. Рассмотрим возможные комбинации, отвлекаясь от лексического значения словоформ и принимая во внимание лишь грамматический вид.

Если личному глаголу совершенного вида в прошедшем времени предшествует другой такой глагол (обозначим эту ситуацию как СВ + СВ), то в тексте мы имеем нарратив, характеризующийся иконическим отражением событий. Семантика этих форм позволяет создать в повествовательном тексте цепь событий, характеризую-

щихся, согласно В. Шмиду [2003], непредсказуемостью, необратимостью, неповторяемостью и консекутивностью. При построении информационной структуры происходит поочередное создание соответствующих признаков. При этом каждый предыдущий сохраняет свою значимость свершившегося факта (ср. *Пришел, увидел, победил*). По-видимому, это верно и для деепричастий совершенного вида и страдательных причастий прошедшего времени.

Вариант НСВ + СВ не столь однозначен. Довольно часто здесь, как и в предыдущем случае, появление глагола совершенного вида в прошедшем времени влечет вытеснение предшествующего признака, как, скажем, в конструкции *говорил долго и внезапно умолк*. Однако это не всегда так, ср. *ехал в автобусе и вдруг подумал, лежал в гамаке и заметил*.

Когда второй признак выражен личным глаголом несовершенного вида прошедшего времени, невозможно однозначно утверждать, имеем ли мы дело с синхронными или сменяющимися друг друга признаками. Можно утверждать, что вариант СВ + НСВ часто связан со сменой типа речи, а именно переходом от сообщения к описанию (примеры см. ниже). Что касается последовательности НСВ + НСВ, она встречается как в дескриптиве, так и в нарративе (*подхожу и спрашиваю*), выражать как синхронные, так и последовательные признаки. Это связано с семантикой форм прошедшего несовершенного, которые могут служить как для описания отдельной ситуации, «наличной к определенному моменту в течении событий» [Маслов 1984: 183], так и для описания регулярно повторявшихся в прошлом ситуаций (в том числе событийного плана). Следовательно, в этих случаях первому глаголу не может быть приписан какой бы то ни было определенный статус.

В теоретическом плане граница между дескриптивом и нарративом не является четкой из-за промежуточного положения, которое занимают, во-первых, динамические описания и, во-вторых, рассказы о рутинных событиях, лишенные осложнения (то, что в англоязычной традиции именуется *habitual narratives*). Однако для нашей задачи однозначное отнесение к той или иной категории несущественно; важно то, какие построения в информационной структуре влечет тот или иной тип текста. Обратимся к примерам.

Спорный статус динамического описания обусловлен тем, что в нем используются глаголы активного действия, но они служат не для рассказа о реальных действиях, совершаемых во времени, а для

характеристики лица, предмета и др. Динамическое описание, как правило, имеет предметную рему, ср.:

*Окно моей комнаты выходило в одичавший сад, заросший смородиной, малиной, лопухом и крапивой вдоль забора. По утрам за окном **возились** воробьи, тучами **налетали** дрозды клевать смородину, доктор не **гонял** их и ягоду не **собирал**. На забор иногда **взлетали** соседские куры с петухом. Петух громогласно **пел**, **вытягивая** кверху шею, **дрожал** хвостом и с любопытством **смотрел** в сад. Наконец он не **выдерживал**, **слетал** вниз, за ним **слетали** куры и поспешно **начинали рваться** возле смородиновых кустов. Еще в сад **забредали** коты и, затаясь возле лопухов, **следили** за воробьями (Ю. Казаков. Арктур — гончий пес).*

Собственно говоря, с точки зрения построения информационной структуры по тексту разграничение между статическим и динамическим описанием несущественно: и то, и другое предполагает приписывание признаков объекту. Более значимым является рассмотренное выше различие между качественной и предметной ремой, связанное с числом создаваемых объектов. Если рема предметная, требуется создавать более одного объекта и, как указывалось выше, встает задача выявления имплицитных связей между ними в пределах одного «суперобъекта».

Нарративы без осложнения, повествующие о регулярно повторяющихся ситуациях, широко распространены в художественной литературе. Поскольку они основываются на использовании глаголов несовершенного вида, которые могут служить и для описания отдельной ситуации, возникает непростая задача автоматической дифференциации нарратива без осложнения и «обычного» повествования.

В отличие от описания (дескриптива), в повествовании (нарративе) нередки вводные слова со значением оценки, а также обстоятельственные слова со значением временной последовательности, призванные отражать динамику повествования [Нечаева 1974: 128–132].

* * *

Отличительная черта рассуждения как типа речи — насыщенность текста субъективной модальностью, что особенно характерно для публицистического стиля. Именно здесь наиболее остро стоит уже неоднократно упоминавшаяся проблема разграничения факта

и мнения. Создаваемые в информационной структуре пропозиции, отражающие диктум (по Ш. Балли), должны снабжаться привязкой к соответствующему модусу и его автору. Это специальная задача, которая будет рассмотрена ниже.

Если для публицистического стиля основная оппозиция типов речи — это рассуждение vs. (описание + сообщение), то для художественного важнее всего разграничивать описание и сообщение. Именно на этой проблеме мы хотим остановиться подробнее. Обратимся к распространенной ситуации, при которой дескриптивные фрагменты (описания персонажей, предметов обстановки, места действия и пр.) вкраплены в нарративный художественный текст. Возникает вопрос: как можно автоматически распознать, где в литературном произведении кончается описание и начинается действие (или наоборот)? Как уже указывалось, от ответа на этот вопрос зависит статус создаваемых в информационной структуре признаков: одновременный (синхронный) или последовательный. Проанализируем, какие языковые средства служат маркерами перехода от нарратива к дескриптиву и обратно⁵².

Пожалуй, самым частотным показателем является смена глагольного вида (если и повествование, и описание строятся в прошедшем времени). В приводимом ниже отрывке можно видеть переход от описания обстановки (глаголов несовершенного вида) к действию (глаголам совершенного вида) и затем к описанию персонажа (снова глаголам несовершенного вида), ср.:

*В горнице **было** тепло, сухо и опрятно: новый золотистый образ в левом углу, под ним покрытый чистой суровой скатертью стол, за столом чисто вымытые лавки; кухонная печь, занимавшая дальний правый угол, ново **белела** мелом; ближе **стояло** нечто вроде тахты, покрытой пегими попонами, упиравшейся отвалом в бок печи; из-за печной заслонки сладко **пахло** щами — разварившейся капустой, говядиной и лавровым листом.*

*Приезжий **сбросил** на лавку шинель и **оказался** еще стройнее в одном мундире и в сапогах, потом **снял** перчатки и картуз и с усталым видом **провел** бледной худой рукой по голове — седые волосы его с начесами на висках к углам глаз слегка **курчавились**, красивое удлиненное*

⁵² Заметим, что графический признак (начало нового абзаца) ненадежен: в общем случае он не является ни необходимым, ни достаточным условием смены типа речи.

лицо с темными глазами **хранило** кое-где мелкие следы осы (И. А. Бунин. Темные аллеи).

Переход от описания к действию могут сигнализировать такие лексические средства, как *однажды, случай, случилось, неожиданно* и пр. Так, длинный абзац, посвященный наружности и характеру Сильвио в начале пушкинской повести «Выстрел», заканчивается фразой *Нечаянный случай всех нас изумил*, — а следующий абзац, в котором речь идет о происшествии за карточным столом, открывается предложением *Однажды человек десять наших офицеров обедали у Сильвио*. В «Капитанской дочке» окончание описания детства главного героя обозначается фразами *Тут судьба моя переменилась. Однажды...* Обратный переход от нарратива к описанию не имеет регулярных лексических средств выражения и обычно маркируется только сменой глагольного вида.

Кроме смены глагольного вида, в художественной литературе также широко представлено переключение с прошедшего времени повествования на настоящее время описания (и обратно). В частности, этот прием регулярно используется для характеристики географической местности:

*Соден — небольшой городок в получасовом расстоянии от Франкфурта. Он **лежит** в красивой местности, на отрогах Таунуса, и известен у нас в России своими водами, будто бы полезными для людей с слабой грудью. Франкфуртцы **ездят** туда большие для развлечения, так как Соден **обладает** прекрасным парком и разными «виртшафтами», где можно пить пиво и кофе в тени высоких лип и кленов. Дорога от Франкфурта до Содена **идет** по правому берегу Майна и вся обсажена фруктовыми деревьями.*

*Пока карета тихонько катилась по отличному шоссе, Санин украдкой **наблюдал** за тем, как Джемма **обращалась** со своим женихом: он в первый раз **видел** их обоих вместе (И. С. Тургенев. Вешние воды).*

При описании персонажа подобное переключение тоже встречается, но значительно реже. В связи с тем, что интересующий нас фрагмент достаточно длинный, приводим его с сокращениями:

*Оборачиваюсь: Грушницкий! Мы **обнялись**. Я **познакомился** с ним в действующем отряде. Он **был ранен** пулей в ногу и **поехал** на воды с неделю прежде меня.*

*Грушницкий — юнкер. Он только год в службе, **носит**, по особенному роду франтовства, толстую солдатскую шинель. У него георгиев-*

ский солдатский крестик. Он хорошо сложен, смугл и черноволос; ему на вид можно дать двадцать пять лет, хотя ему едва ли двадцать один год. Он **закидывает** голову назад, когда говорит, и поминутно **крутит** усы левой рукой, ибо правой **опирается** на костыль. **Говорит** он скоро и вычурно <...>

Впрочем, в те минуты, когда **сбрасывает** трагическую мантию, Грушницкий довольно мил и забавен. Мне любопытно видеть его с женщинами: тут-то он, я думаю, **стареется!**

Мы **встретились** старыми приятелями. Я **начал** его расспрашивать об образе жизни на водах и о примечательных лицах (М. Ю. Лермонтов. Герой нашего времени).

Заметим, что схожий механизм может использоваться также при переходе от нарратива к рассуждению и обратно, ср.:

Он на пороге **остановился**: ему **хотелось** пожать мне руку... И если б я показал ему малейшее на это желание, то он бросился бы мне на шею; но я **остался** холоден, как камень — и он **вышел**.

Вот люди! все они таковы: **знают** заранее все дурные стороны поступка, **помогают**, **советуют**, даже **одобряют** его, видя невозможность другого средства, — а потом **умывают** руки и **отворачиваются** с негодованием от того, кто имел смелость взять на себя всю тяжесть ответственности. Все они таковы, даже самые добрые, самые умные!..

На другой день утром, получив приказание от высшего начальства отправиться в крепость N, я **зашел** к княгине проститься (М. Ю. Лермонтов. Герой нашего времени).

Однако в целом этот признак ненадежный, так как в пределах одного нарративного фрагмента возможно сочетание форм прошедшего времени с настоящим или будущим в стилистических целях — для оживления излагаемых событий, для подчеркивания длительности действия и пр.

Бывает, что переходы от нарратива к дескриптиву и назад не маркируются ни сменой глагольного вида, ни времени. Что же в таком случае выступает носителем соответствующей информации? Как читающий текст человек понимает, что такой переход имеет место? Есть примеры, в которых соответствующим сигналом является употребление глагола с иными семантическими характеристиками, чем предшествующие глаголы. Так, в примере ниже глаголы движения *идти* и *кланяться* употребляются в нарративном фрагменте, а статический глагол *быть* — при описании внешности персонажа, ср.:

Она **шла** босая, с дерюжной сумой через плечо, подпираясь высокой палкой, и в деревьях и селах молча **кланялась** перед каждой избой. Мальчик **шел** за ней сзади, тоже с мешком через плечико, в старых башимаках ее, разбитых и затвердевших, как те опорки, что валяются где-нибудь в овраге.

Он **был** урод. У него **было** большое, плоское темя в кабаньей красной шерстке, носик расплющенный, с широкими ноздрями, глазки ореховые и очень блестящие. Но когда он улыбался, он **был** очень мил (И. А. Бунин. Дурочка).

Однако это наблюдение не следует возводить в общее правило. Выше уже упоминалось о том, что в художественном тексте одни и те же глаголы могут использоваться как в нарративе, так и в дескриптиве⁵³. Показателен, например, следующий фрагмент, где пропозиция с глаголом *вливался* завершает нарратив о продвижении судна, а последующий текст посвящен описанию, но также преимущественно посредством глаголов действия. Ср.:

Пароход медленно **вливал** в Азию — в приторный дым кизяка, в туманы пыли, в огни прибрежных костров, в кочевья, раскинутые вокруг города, в заунывные **крики** киргизов, **тянувших** на песке мокрые невода.

Шелудивые псы **бежали** по берегам с хриплым лаем, за ними **мчались** голые киргизские дети. Пестрое тряпье, развешанное на шестах, **приветствовало** нас подобно знаменам Тамерлановых полчищ. **Рыдали**, как двери на тугих пружинах, облезлые верблюды.

По берегам **брели** киргизы, **волоча** хвостами по пыли **трепещущих** сазанов и судаков. По их грязной чешуе стекали чистые капли крови.

Чигири **скрежетали** во внезапной темноте, похожей на кофейную гущу. Огненная надпись «Эмбанефть», висящая в воздухе, как бы **открывала** вход в тяжелую пустыню (К. Г. Паустовский. Подводные ветры).

Разумеется, можно трактовать три последних абзаца как нарратив (или скорее цепь из нескольких нарративов: про псов, верблюдов, киргизов и пр.), но тогда встает вопрос о соотносительности между

⁵³ Ср.: «Факты свидетельствуют о том, что семантика глаголов, указывающих на состояние предмета или его действие, движение, не соответствует понятиям статичного в описании и динамичного в повествовании. Глагол со значением действия, даже движения, может передавать статичность содержания в описании, а глагол состояния — динамику повествования» [Нечаева 1974: 114].

собой этих нарративов, в том числе отношениях вложенности. Заметим, что приведенный отрывок не является каким-то редким, исключительным или присущим авторскому стилю конкретного писателя. Вновь обратившись к творчеству Бунина, можно найти схожий пример, в котором глаголы действия также используются для описания обстановки. Переход к нарративу маркируется здесь словом *вдруг* (аналогичную функцию могли бы также выполнять обороты и *тут*, в *это мгновение* и т. п.):

В глаза бьет зарево пожара вдали направо. Тележка стоит против того леска, что открывался при зарницах. Лесок от зарева стал теперь черным и весь зыбко дрожит, как дрожит и все поле перед ним в сумрачно-красном трепете от того жадно несущегося в небе пламени, которое, несмотря на даль, полыхает с бегущими в нем тенями дыма точно в версте от тележки, разъяряется все жарче и грознее, охватывает горизонт все выше и шире, — кажется, что жар его уже доходит до лица, до рук, виден даже над чернотой земли красный переплет какой-то сгоревшей крыши. А под стеной леса стоят, багрово серея, три больших волка, и в глазах у них мелькает то сквозной зеленый блеск, то красный, — прозрачный и яркий, как горячий сироп варенья из красной смородины. И лошади, шумно всхрапнув, вдруг диким галопом ударяют вбок, влево, по пашне, малый, на вожжах, валится назад, а тележка, со стуком и треском, мотаясь, бьется по взметам... (И. А. Бунин. Волки).

Напоследок заметим, что в художественной литературе повествование и описание иногда столь тесно переплетены, что граница между ними идет по отдельным предложениям внутри абзаца (если не по частям одного предложения). В таких случаях говорят о «контаминации функционально-смысловых типов речи» [Нечаева 1974: 195–203], или о «сложных смешанных типах речи» [Валгина 2003: 94], ср.:

Я пошел вправо через кусты [повествование]. Между тем ночь приближалась и росла, как грозовая туча; казалось, вместе с вечерними парами отовсюду поднималась и даже с вышины лилась темнота [описание]. Мне попала какая-то неторная, заросшая дорожка; я отправился по ней, внимательно поглядывая вперед [повествование]. Все кругом чернело и утихало, одни перепела изредка кричали [описание]. Небольшая ночная птица, неслышно и низко мчавшаяся на своих мягких крыльях, почти наткнулась на меня и пугливо нырнула в сторону. Я вышел на опушку кустов и побрел по полю межой

[повествование]. *Уже с трудом различал я отдаленные предметы; поле неясно белело вокруг; за ним, с каждым мгновением надвигаясь громадными клубами, вздымался угрюмый мрак. Глухо отдавались мои шаги в застывающем воздухе. Побледневшее небо стало опять синеть — но то уже была синева ночи. Звездочки замелькали, зашевелились на нем* [описание] (И. С. Тургенев. Бежин луг).

* * *

Проблема разграничения повествования и описания встает не только при анализе художественной литературы. Выше мы рассматривали ситуацию, где в нарративный текст вставлены дескриптивные фрагменты. Для полноты картины обратимся теперь к противоположному случаю: в дескриптивный текст включен нарративный фрагмент. Это широко встречается, например, в путеводителях. Нарратив здесь либо служит для рассказа об истории возникновения достопримечательности (в этом случае наблюдается переключение временных форм глагола), либо обозначает типовой маршрут путешественника. Приведем примеры обоих случаев.

*Костел Св. Микулаша **создан** двумя выдающимися пражскими зодчими — Кристофом и Килианом Игнацом Динценхоферами. Отец, Кристоф, **построил** в начале XVIII столетия неф и западный фасад, а сын спустя 40 лет хор и купол. **Завершил** постройку созданием колокольни (1755) Ансельмо Лураго. Размеры костела **впечатляют**: его куполу на высоте 75 метров могла бы позавидовать смотровая площадка на холме Петршин. Не **имеют** себе равных и росписи сводов главного нефа «Апофеоз Св. Микулаша» (Я. Л. Кракер), занимающие площадь 1500 м². В интерьере **обращают** на себя внимание живопись К. Шкреты и скульптура И. Ф. Платцера⁵⁴.*

Если нарративный фрагмент посвящен изложению маршрута путешествия, он может быть оформлен в настоящем времени, как и дескриптивный текст. Переход от статического описания достопримечательности к (воображаемому) движению можно распознать по широкому использованию лексики движения, а также частым отсылкам к эгоцентрической системе пространственных координат, ср.:

Проилеи — парадный вход на Акрополь, а само слово можно примерно перевести как «преддверие». Второе по времени постройки здание

⁵⁴ Прага: путеводитель (серия Polyglott). СПб., 1996. С. 56.

на Акрополе после Парфенона сооружено по проекту Мнесикла в 437–432 годах до н. э. Задача здания — стать архитектурным предисловием к Парфенону, мощью и гармонией создавать торжественное настроение поднимающемуся на Акрополь. С этой задачей постройка Мнесикла, несмотря на разрушения, великолепно справляется до сих пор.

Посетитель **подходит** к Пропилям сбоку и внезапно **оказывается** на широкой мраморной лестнице, охваченной двумя крыльями-портиками. По мере **приближения** колоннады крыльев **расступаются** и уступают все поле зрения центральному фасаду из дорических колонн, которые, в свою очередь, **раскрываются** посередине и уводят внутрь здания, в коридор, обрамленный ионическими колоннами. Коридор **приводит** к сверхчеловеческих размеров центральной двери, сразу за которой — внутренний фасад из дорических колонн. **Пройдя** под колоннами, **оказываетесь** на Священной дороге с парадным видом на Парфенон. Рядом с центральной «дверью для богов» — еще четыре, уменьшающиеся от центра к краям, для людей.

Левое крыло Пропилей занимала Пинакотекa — картинная галерея. Это одно из первых известных нам собраний живописи⁵⁵.

Изложение маршрута может маркироваться и другими средствами, в числе которых следует упомянуть переключение с нейтрального, «безличного», рассказа на персонализированное мы-или вы-повествование, а также использование директивных речевых актов, ср.:

Отдохнув, **спустимся** к Висле, **полюбимся** ее берегами с обзорной площадки, а дальше **направимся** к городским воротам — барбакану⁵⁶.

В Новом городе **вы повстречаете** огромное количество костелов: костел Святого Духа, костел Святого Казимира и многие другие⁵⁷.

Зайдите в новое здание библиотеки Варшавского университета, которая находится по адресу...⁵⁸

Показателен следующий отрывок, в котором можно наблюдать чередование основных содержательных компонентов туристического дискурса — нейтральной дескрипции (в настоящем времени),

⁵⁵ Греция: путеводитель (серия «Вокруг света»). М., 2011. С. 56–57.

⁵⁶ Польша: путеводитель (серия ORANGEвый гид). М., 2013. С. 76.

⁵⁷ Там же. С. 78.

⁵⁸ Там же. С. 85.

нарратива об истории создания памятника (в прошедшем времени) и актуального нарратива-маршрута (снова в настоящем):

В юго-восточной части Рыночной площади (Rynek Główny) находится небольшой храм. Это костел Святого Войцеха — первого святого Польши. Он был построен на месте деревянного храма в XII в. Интересно, что костел находится на 2,5 м ниже уровня земли.

Далее мы идем к одиноко возвышающейся Ратушной башне. Башня — это то, что осталось от готической ратуши XIV в.⁵⁹

РАЗГРАНИЧЕНИЕ ДИКТУМА И МОДУСА

Для дальнейшего изложения методики коммуникативного анализа текста следует обсудить комплекс вопросов, связанных с оппозицией «диктум — модус», а именно: разграничение пропозиций факта и мнения, определение субъекта мнения, анализ оценочного компонента суждений и его отражение в информационной структуре текста, учет коннотаций лексических единиц и др.

Содержание языковых средств, используемых автором для воплощения сведений об известных ему фактах, своей трактовке событий, отношении к ситуации, о качестве взаимоотношений между людьми и т. д., состоит из двух различных компонентов: диктумной и модусной части⁶⁰. Каждый из этих понятийно-семантических компонентов содержания языковых средств характеризует принципиально разные стороны отраженной действительности.

Диктумная часть (или диктум) — это верифицируемый (рационально-когнитивный) компонент языковых средств, который составляют семантические признаки, характеризующие объективное, конкретное, предметное знание, основанное на эмпирическом, ра-

⁵⁹ Польша: путеводитель. С. 128.

⁶⁰ В отличие от теории высказывания Ш. Балли, в которой термины средневековой схоластики «диктум» (лат. *dictum* — слово, выражение) и «модус» (лат. *modus* — способ) используются для обозначения объективной и субъективной частей суждения (предложения), в данной методике этими терминами обозначаются частные компоненты семантической функции различных видов языковых средств (морфологических, синтаксических, лексико-семантических, функционально-прагматических). Информация о рациональном содержании языкового средства именуется диктумом, а свойственное определенной части языковых средств выражение субъективно-личного отношения к рациональному содержанию — модусом.

циональном освоении результатов непосредственного воздействия на органы чувств человека вещных, материальных свойств предмета. Диктумом охватываются материально-эмпирические, условно-фактические признаки понятий и утверждений, отражающие в сознании объективные свойства и качества денотатов (т.е. внеязыковых референтов, понятия о которых обладают категорией вещиности). Сюда входят понятия, обобщающие чувственные образы воздействия внешней силы на материально-вещественные объекты, включая человека, понятия, характеризующие природные процессы, конкретные физические действия человека. Диктумные признаки понятий связываются в суждениях на основании природной смежности, совместимости, равносущности их реальных денотатов на оси «целое — часть» фрагмента действительности, конкретизируя «часть» реальности. Диктумная часть содержания языковых средств осознается как фактическое (констатирующее, объективное) знание о материально-вещественных объектах, об их количественных характеристиках (по параметру «мало — много»), внешних объективно-качественных свойствах и предикатных способах их пространственной и временной реализации (по параметрам «далеко — близко», «быстро — медленно», «долго — кратко», «часто — редко», «начало — конец»).

Модусная часть (или модус) — это интерпретирующий (эмоционально-прагматический) компонент языковых средств, который включает семантические и стилистические признаки, характеризующие субъективно-качественные и экспрессивно-оценочные аспекты значения высказывания. Модусную часть содержания составляют значения понятийных признаков, отражающих субъективно-ментальную, оценочную характеристику объектов высказывания. С помощью модуса характеризуется степень значимости (оценка важности, ценности, полезности), целесообразности или необходимости предмета речи для говорящего. Модус объединяет эмоционально-экспрессивную, регулятивную часть признаков понятий и суждений, которые связаны субъективными конструктивно-логическими обобщениями.

В модусной части фиксируется опыт оценочной интерпретации качественных (волевых, интеллектуальных, моральных, нравственных, эмоциональных, идеологических, этических) сторон деятельности одушевленных существ, стихийных проявлений сил природы и природных процессов. Модусные части суждений с признаками субъективно-качественной характеристики свойств объектов и про-

явлений предикатных признаков образуют мыслительные структуры, отражающие результаты личного восприятия качественных характеристик материально-вещественных объектов.

Модусы выражают оценки лица, его деятельности, а также оценки поступков физического лица с точки зрения здравого смысла, морали или с правовой точки зрения (в той мере, в которой это может понимать любой дееспособный гражданин, не имеющий специальных познаний в области юриспруденции). Такие понятийные признаки связаны имплицативными отношениями, и в суждениях о денотатах осознаются как выводы из фрагмента действительности на оси «общее — частное», наслаиваясь при этом на диктумную часть значения как «частное».

Модусная часть содержания языковых средств оказывает регулятивное речевое воздействие на мнение и сознание слушателей, поскольку обладает свойствами субъективной информации (с параметрами «плохой — хороший», «справедливый — несправедливый», «честный — нечестный», «объективный — субъективный» и т.п.). В частности, укажем на значения с негативным модусом, интерпретирующим диктумную часть лексического содержания языкового знака: *сговор, шельмовать, клеветать, всучить, навязать, самоуправство, пресловутый, вопиющий, зачинщик, пособник, подлец, фанатик, экстремист, военизна, бодрячество, потуги, интервенция, сборище, произвол, водяра, винище, морда* — или на модусную часть значения словосочетаний: *нелепые бредни, подлая демагогия, пустые фантазии, тоталитарная секта, беспробудный пьяница, усатая гадина* и т.п.

Продолжая иллюстрации диктумных и модусных языковых средств, процитируем слова П. А. Столыпина: *«Народ, не имеющий национального самосознания, — есть навоз, на котором произрастают другие народы»*. Здесь используются значения слов с диктумным компонентом, которые участвуют в конструкции образного сравнения (*народ есть навоз*). Само сравнение, правда, содержит негативную характеристику субъекта сравнения. Приведем также высказывание А. В. Суворова: *«Я русский, — какой восторг!»*, состоящее из диктумной (*я русский*) и модусной (*какой восторг!*) части⁶¹.

⁶¹ В последнее время обе цитаты используются в национально ориентированном дискурсе. См., например: URL: http://www.soprotivlenie.su/slovo/ya_russkiy_kakoy_vostorg.html (дата обращения: 15.03.2018); URL: <http://slavyanskiysoyuz.ru/26-my-russkie-kakoy-vostorg.html> (дата обращения: 15.03.2018); URL: [http://gidepark.ru/user/875461648/content/653617\\$](http://gidepark.ru/user/875461648/content/653617$) (дата обращения: 15.03.2018); URL: <http://trusia.livejournal.com/5980.html> (дата обращения: 15.03.2018) и мн. др.

Выбор языковых средств с определенным диктумом и модусом осуществляется автором и зависит от качеств темы сообщения и отношения автора к предмету речевой характеристики. Диктумный аспект языковых средств обеспечивает указание на объективные признаки предмета речи, характеризующие внеязыковой референт, поэтому общность диктумных компонентов определяется главным объектом речи и темой сообщения. Языковые средства с нейтральным модусом широко используются в сообщениях о ситуациях, событиях, явлениях, о лице, поступках людей, которые обычно не описываются в ценностных категориях (к такой информации неприменимы оценки «хорошо» и «плохо»).

Если диктумом обладают все инструментальные и коммуникативные языковые средства, то модусом (положительным или отрицательным) обладают далеко не все из них. Именно значения семантических признаков диктума создают когеренцию языковых средств, а вот использование в речи средств с модусным компонентом целиком зависит от воли автора и не влияет на порядок и правила понятийно-семантического связывания частей высказывания. Увеличивая количество языковых средств, обладающих модусом, автор усиливает речевое воздействие на эмоциональную сторону сознания слушателей, поскольку модусный аспект языковых инструментов обеспечивает выражение авторского отношения к предмету мысли и его признакам. Все модусные коннотации должны обязательно указываться в словарном описании значений слов, которое является ядром используемых нами процедур коммуникативно-понятийного анализа текста.