

# Выражение обусловленности в естественном языке<sup>1</sup>

Л.В. Савинич, В.Л. Стефанюк

**Аннотация.** Изучается новая для области машинной обработки текстов проблема выделения в тексте элементов, которые позволяют чисто формальными методами и приёмами представить текстовое пространство регламентирующего характера в виде строгих логических правил, выражающих связи между обсуждаемыми фактами и явлениями. Вводится понятие обусловленности как обобщение понятия причинности и изучаются соответствующие языковые средства.

## Введение

### Постановка задачи

В статье формулируется новая задача выделения в тексте элементов, позволяющих чисто формальными методами преобразовать текст в совокупность строгих логических правил, типичных для интеллектуальных систем. Для решения этой задачи вводится понятие обусловленности как обобщение понятия причинности. В результате можно надеяться, что содержание текста может быть представлено как совокупность экспертных правил и заключений рекомендательного характера. Такая постановка задачи, очевидно, отличается от проблемы «смысл ↔ текст», снимаемая многими трудностями, отмеченные в [1], и заменяя эту важную проблему задачей определенного перефразирования текста, делающего его пригодным для использования в экспертных системах.

Если говорить о содержании текста, естественно выделить две основные компоненты – логическую (А) и внелогическую (В). Разумеется, в обычном тексте эти компоненты могут

произвольным образом пересекаться и сосуществовать, тогда как в специализированных текстах – инструкциях, учебниках, научных статьях – логическая часть может быть специально усилена.

Логическая часть текста посвящена изложению логических взаимосвязей содержательных элементов того материала, который в тексте имеется. Эта часть, в свою очередь, может быть двух видов:

А1 – внутренняя логика - логическая компонента, которая объясняет логику организации текста, позицию автора и т.д. Она нужна авторам, чтобы построить сам текст, нам, например, чтобы выстроить план настоящей публикации. Поэтому её можно назвать также рабочей логикой текста.

А2 - внешняя логика, или, точнее, общезначимая утилитарная логика. Это логика тех сведений, которые автор хочет сообщить читателю: что именно тот, по мнению автора, должен усвоить и принять для себя в будущей своей деятельности. При этом материал подаётся так, чтобы не противоречить характеру и порядку мыслительных рассуждений читателя, чтобы

<sup>1</sup> Работа частично финансировалась РФФИ по проектам 05-01-00382 и 07-07-00391, а также по программе №14 Президиума РАН. Предварительный вариант работы был доложен на конференции САИТ-2007.

быть для него убедительным. Типичным примером здесь является литература учебного характера, научного, правового и т.д.

Упомянутая выше внелогическая компонента текста (В) призвана передать такие элементы, как эмоциональную окрашенность текста, его эстетическое содержание, стилевую принадлежность, выразительность языка и т.п.

Задача этой компоненты, конечно, в совокупности с логической компонентой – заинтересовать читателя, заинтриговать его, чтобы тот приступил к чтению статьи или книги и увлёкся ею. В принципе, текст может и не содержать общезначимой логической компоненты, т.е. компоненты А2. Например, чисто детективная развлекательная история обычно опирается на логику происходящих событий, но эта логика нужна автору только для поддержания интереса читателя к описываемым событиям, ничего "полезного" ему не сообщая.

Поэтическая форма - это другой пример внелогической, эстетической компоненты. Здесь на читателя воздействует сама форма подачи материала: размер, ритм, рифмы, мелодика, аллитерация и т.п. Известно, что бывают (или были) бессодержательные стихи. Но настоящее творчество, конечно, связано с созданием чего-то поучительного, напутственного для читателя. Только тогда стихи признаются талантливыми.

Например, строка А.Плещеева "ласточка с весною в сени к нам летит", на первый взгляд, не содержит отдельной логической компоненты общезначимого характера. Хотя для юного читателя, может быть, важно будет узнать, что "Если наступит ВЕСНА, то ПРИЛЕТИТ, по крайней мере, одна ЛАСТОЧКА".

Нужно отметить, что мы пока рассматриваем текст изолированно от той культурной среды, в которой этот текст создаётся. Наличие такой среды позволяет писателю прибегать к использованию аналогий, намеков и т.д. Но ниже мы будем стараться рассматривать изолированный текст сам по себе.

На самом деле, давая ссылки на работы других авторов, мы все же помещаем нашу статью в определённую культурную среду, но делаем это наглядно, явным для читателя образом,

стараясь избегать вводных фраз типа "хорошо известно, например, что ...".

В заключение надо сказать, что обе компоненты являются важными и подвергаются разнообразному анализу в литературоведении и т.п. Многие вопросы здесь далеки от их окончательного решения, но нас заинтересовало исследование способов выражения общезначимых логических утверждений, для которых в языке наработано достаточно много средств. Выделению и изучению этих средств посвящена основная часть нашей статьи.

Интерес к этому направлению исследования обусловлен бурным развитием искусственного интеллекта, в котором, начиная с работ Т.Винограда и И.Мельчука [2, 1], формальному анализу фраз естественного языка уделяется огромное внимание. Эта тема становится особенно актуальной с появлением экспертных систем, в которых центральное место занимает вопрос возможности извлечения знаний из текста с помощью автоматических или машинных процедур [3].

Настоящая статья является развитием работы [4] и ставит своей целью обозначить возможные способы выявления семантических структур текста со значением обусловленности. Предполагается, что следующий этап будет посвящён использованию представленных конструкций в экспертных системах. Ожидается, что трансформация любого или только заданного типа дискурса в конструкции с отношениями обусловленности позволит осуществлять поиск требуемых конструкций и их значений при помощи компьютера [5].

В этой связи весь анализ подкреплялся статистическими данными, собранными на конкретном образце делового текста, так как количественные показатели могут свидетельствовать и о типе дискурса.

В данном исследовании отдаётся предпочтение понятию *обусловленность*, которое обобщает принятый в различных областях знаний концепт *причинность*, описывающий причинно-следственные связи.<sup>2</sup> Сущность понятия *причинность* – "порождение причиной следствия" [7, с.1074]. В этом аспекте *обусловлен-*

<sup>2</sup> Лингвистический анализ смыслов 'причина' и 'цель' в естественном языке [6].

ность характеризует тип отношений между несколькими действующими лицами (агентами), ситуациями, явлениями действительности, включая в себя и условие совершения действия. Поэтому на наш взгляд понятие *обусловленность* является более емким.

Обусловленность, или каузальность, т.е. причинность в широком смысле слова, объединяет в себе существенно более богатый спектр значений, таких, как предпосылка, основание, обоснование, подтверждение, доказательство, аргумент, довод, предопределённость, посылка, повод, предлог, стимул, целевая мотивировка. Весь этот круг отношений предполагает такую связь ситуаций, при которой одна служит основанием для реализации другой [8, с.562]. Кроме того, причинность предполагает связь, происходящую *во времени*, тогда как свойство обусловленности временной последовательности очень часто не предполагает. В анализируемом материале нас будет интересовать также и интуитивный аспект изложения, то есть аспект, который автором подразумевается.

## 1. Способы выражения отношений обусловленности

Наиболее очевидными в номенклатуре средств, оформляющих структуры с отношениями обусловленности, являются причинно-

следственные союзы. На синтаксическом уровне эти союзы являются маркерами отношений обусловленности. Однако данную функцию могут выполнять и другие служебные части речи (предлоги, предложные сочетания, частицы), которые будут приведены ниже.

Для более полного представления конструкций со значением обусловленности на лексическом уровне анализировались классы знаменательных частей речи (существительных, прилагательных, причастий), составляющих данные конструкции, и фразеология, характерная для данного типа дискурса. Отдельно были выделены каузативные глаголы, так как причинно-следственные отношения уже входят в семантику этих глаголов.

На синтаксическом уровне отмечена возможная логическая трансформация придаточных предложений в конструкции со значением обусловленности.

И наконец, отношения обусловленности выделялись при логическом анализе смыслового содержания текста – актуальном членении предложения.

## 2. Использование союзов

Как было отмечено выше, для выделения формальных показателей – маркеров, фиксирующих отношения обусловленности в конст-

Союзы, выражающие условные отношения	Союзы, выражающие причинные отношения
<i>если, если бы, в случае если, в том случае если, на случай если, при условии что, при том условии если, при том условии, что;</i>	<i>потому что, потому как, потому, так как, постольку, раз, ибо, ведь, оттого что, из-за того что, ради того что, благодаря тому что, затем что, через то что, для того что, благо, как;</i>
стилистически окрашенные союзы: <i>ежели, ежели бы, когда бы, коли (коль), коли бы, добро бы, кабы, как бы, коль скоро, буде, раз.</i>	книжные и официальные: <i>вследствие того что, в результате того что, в силу того что, на основании того что, исходя из того что, по причине того что, по той причине что, в связи с тем что, ввиду того что, под видом того что, под предлогом того что, под тем предлогом что, тем более что;</i> стилистически окрашенные союзы: <i>тем паче что.</i> <sup>3</sup>

<sup>3</sup> Сопоставительный семантический анализ условных и причинных союзов приводится, например, в [9].

рукциях разного типа дискурса, – были представлены характерные и общеизвестные на синтаксическом уровне союзы со значением обусловленности. Академическая грамматика [8] выделяет две группы таких союзов:

В анализируемом нами тексте официально-делового дискурса используются только три условных союза, оформляющих причинно-следственные отношения: *если; в случае; в случае / -ях, когда*, причем представлены они в незначительном количестве. Так, при общем объёме текста в 3261 слово на долю условных союзов приходится 30 употреблений.

### 3. Субституты союзов со значением обусловленности

Логический анализ текста показал, что, при определённой логической трансформации, отношения обусловленности могут выражать также конструкции с другими классами служебных слов – предлогами, предложными сочетаниями и частицей *ли*. Таких компонентов в синтаксической структуре предложения оказалось значительно больше по сравнению с условными союзами. Среди заместителей условных союзов в анализируемом тексте в первую очередь выделяются сочетания:

с предлогом *с*, а также сложными предлогами, в состав которых входит этот предлог (*с наступлением указанного срока* = 'когда / если наступает указанный срок'; *с участием иностранных граждан* = 'если участвуют иностранные граждане'; *с намерением причинить* = 'если намерен причинить'; *с нарушением требований* = 'если нарушены требования'; *наряду с другими недостатками* = 'если имеются другие недостатки');

с предлогом *по* (*по достижении восемнадцатилетнего возраста* = 'когда / если достигнет восемнадцатилетнего возраста'; *по истечении* = 'когда / если истечёт ...');

с предлогом *без* (*гражданин, осуществляющий предпринимательскую деятельность без образования юридического лица* = 'если не образовано юридическое лицо');

с предлогом *за* (*за причинённый водителем ущерб* = 'если причинён водителем ущерб');

с предлогом *при* (*при наличии* = 'если имеется в наличии / наличествует');

со сложным предлогом *в соответствии с* (*в соответствии с законом / пунктом / этим* = 'если соответствует закону / пункту / этому');

с частицей *ли* (*...в зависимости от того, осуществлялись ли эти права разумно и добросовестно, ...* = 'если эти права осуществлялись разумно и добросовестно'; *...независимо от того, уполномочен ли он вести дела товарищества* = ..., 'если уполномочен он вести дела товарищества').

Как видим, большинство представленных предлогов и предложных сочетаний адекватно передают смысл условного союза *если*. Другие же конструкции, несмотря на дополнительные смысловые оттенки [временные], также сохраняют значение условности.

Практически во всех приведённых примерах предлоги, осуществляя свои синтаксические функции, управляют отглагольными существительными. На этом основании можно заключить, что определённая логическая трансформация предложных словосочетаний возможна с опорой на значения глагольности.

## 4. Лексический анализ текста

Отражая тем или иным образом логическую структуру высказывания, синтаксические отношения образуют в языке формальные признаки слов, обозначающих предмет высказывания, и противопоставленные им признаки слов, типично обозначающих предикацию. С целью представления для поисковых систем формальных признаков слов, с одной стороны, и выявления понятийного аппарата текста, с другой стороны, был предпринят анализ лексического состава текста, в результате чего определены характерные для данного типа дискурса классы лексических единиц и форманты, маркирующие их.

Поскольку значение причинно-следственных связей уже входит в семантику некоторых глаголов, то лексический анализ текста был начат с глаголов.

### 4.1. Каузативные глаголы

Многие из входящих в анализируемый текст глаголов относятся к группе каузативов: *руководить, причинить, отказать, удовлетворить / -яются, распоряжаться, признаёт / -ся, освобождается, регулирует / -ют / -ются*. Для

большей ясности можно привести толкование этого широкоупотребительного в лингвистике термина: «каузатив – глагол со значением 'делать так, что некая ситуация начинает иметь место'» [10, с. 46], иными словами: одно действие обуславливает другое действие.

Давая характеристику этим глаголам, специалисты отмечают, что большую роль в формировании каузативных отношений играет фактор разграничения действующих лиц: инициатора действия и его исполнителя, соответственно, фиксирование двух событий – акта побуждения и акта осуществления [11, с.173], т.е. каузирующее событие и каузируемое событие. Другой важный фактор для формирования каузативных отношений – синтетическое или расщеплённое обозначение таких событий, впервые описанное и проиллюстрированное в [10, с.153–155].

С точки зрения участия в каузативных отношениях лиц и событий были выделены несколько типов глаголов, в семантику которых включён компонент каузативности [11, с.176].

В анализируемом нами тексте выделенные каузативные глаголы относятся к разным типам предложенной классификации. Например, глагол *руководить* можно отнести к типу "личностных каузативов, представляющих некоторое событие как результат воздействия одного лица на волю другого"; глагол *причинить* – к "событийным каузативам, представляющим событие как результат воздействия некоторой ситуации (личной или безличной) на агента действия или состояния" и т.д.

Причём, в семантике этих глаголов отражается разное обозначение событий, например, в глаголе *руководить* – каузирующее событие, в глаголе *освобождается* – каузируемое и т.д.

Так как рассматриваемый нами текст имеет преимущественно декларативный, регламентирующий характер, все глаголы в них используются соответственно или в форме инфинитива, или безлично (в форме 3 лица единственного числа), или неопределённо-лично (в форме 3 лица множественного числа). Данные словоформы и образующие их форманты могут служить маркерами для поисковых систем.

Всего в контрольном образце текста использовано 63 глагола и в общей сложности насчитывается 216 глагольных употреблений.

#### 4.2. Анализ прилагательных

Прилагательные как часть речи имеют свои известные морфологические и синтаксические характеристики.

Употребляясь атрибутивно (т.е. в качестве определения к существительному), прилагательное в такой синтаксической конструкции находится в контактной препозиции по отношению к определяемому существительному: *интеллектуальная деятельность*. Употребляясь предикативно, прилагательное может соединяться с существительным также через глагол-связку и стоит в постпозиции по отношению к существительному: *Они свободны в установлении своих прав и обязанностей. Гражданин ... может быть признан судом недееспособным...*

В анализируемом нами тексте было выявлено 94 прилагательных и в общей сложности 350 употреблений прилагательных.

Абсолютное большинство прилагательных принадлежит к разряду относительных (т.е. обозначающих свойство предмета через его отношение к другому предмету, признаку, событию): *авторский* (договор), *бытовые* (сделки), *внебюджетные* (фонды), *договорные* (обязательства), *кредитные* (учреждения), *наркотические* (средства), *психические* (расстройства), *спиртные* (напитки).

Употребляясь атрибутивно в препозиции по отношению к существительному, такие прилагательные могут образовывать с определяемым словом устойчивые и терминологические словосочетания, характерные для официально-деловой сферы деятельности. Эти устойчивые словосочетания также могут являться опорными для разнообразных поисковых систем.

#### 4.3. Анализ причастий

Причастие совмещает в себе свойство глагола и прилагательного и, подобно другим классам лексических единиц, обладает набором характерных формальных признаков. Являясь нефинитной формой глагола, причастие обозначает признак, связанный с действием и

употребляемый атрибутивно (т.е. в качестве определения). Полные формы причастий могут употребляться как в препозитивной, так и в постпозитивной синтаксической позиции. Однако, входя в определительную конструкцию с зависимыми словами, полные причастия могут быть обособленными (обособленный причастный оборот). Такие конструкции, как правило, находятся в контактной постпозиции по отношению к определяемому слову и близки по значению определительным, условным и другим типам придаточных предложений. Например: *...лицами, осуществляющими предпринимательскую деятельность, ...* = 'лицами, если осуществляют предпринимательскую деятельность'; *...отношения, основанные на равенстве, ...* = 'отношения, если основаны на равенстве'; *...акты, содержащие нормы гражданского права, ...* = 'акты, если содержат нормы гражданского права'.

Помимо имеющихся у всех причастий определённых морфологических показателей, дополнительным сигналом или знаком для поисковых систем может служить и обязательное пунктуационное выделение этих синтаксических конструкций.

Краткие формы причастий выступают в роли сказуемого (*...договором не предусмотрено возмещение убытков в меньшем размере...*). В соединении с глаголом *быть* краткие формы образуют аналитическую форму страдательного залога (*...его право не было нарушено...; ...никто не может быть ограничен в правоспособности и дееспособности...*).

В контрольном тексте насчитывается 58 полных причастий и 153 их употреблений, а также 8 кратких причастий и 28 их употреблений.

#### 4.4. Анализ существительных

Существительные представляют самую многочисленную группу лексических единиц (274 существительных и в общей сложности 1323 употребления в тексте). Абсолютное большинство их – существительные отвлечённые, со свойствами им формальными показателями.

Данные существительные составляют основную понятийный костяк текста и вместе с полными прилагательными и причастиями образуют устойчивые конструкции и конструкции

терминологического характера. Однако помимо своей основной номинативной функции (т.е. функции называния), существительные в данном типе дискурса выступают очень часто и в атрибутивной функции, образуя ряды номинативных конструкций: *место жительства гражданина; права авторов произведений науки, литературы и искусства; недопустимость лишения и ограничения правоспособности и дееспособности гражданина; осуществление процедуры признания банкротом* и т.п.

### 5. Придаточные определительные предложения

Аналогично причастным оборотам, логической трансформации в конструкции со значением обусловленности могут подвергаться и отдельные виды придаточных определительных предложений с союзным словом *который*: *При искажении или использовании имени гражданина способами или в форме, которые затрагивают его честь, достоинство или деловую репутацию, применяются правила...* = 'при искажении или использовании имени гражданина способами или в форме, если затрагивают его честь, достоинство или деловую репутацию, применяются правила...'; *Индивидуальный предприниматель, который не в состоянии удовлетворить требования кредиторов, может быть...* = 'индивидуальный предприниматель, если не в состоянии удовлетворить требования кредиторов, может быть...'. Однако подобная логическая трансформация возможна, если союзное слово является субъектом действия в предложении. В позиции объекта союзное слово затрудняет логическую замену.

Подобно причастным оборотам, придаточные определительные предложения также выделяются с обеих сторон запятыми, что в совокупности с союзным словом *который* / *-ые* может служить дополнительным сигналом или знаком для поисковых систем.

### 6. Актуальное членение предложения

Анализ официально-деловых текстов на уровне смысла установил, что в данном типе дискурса преобладает однотипная структура предложений с чётким и упорядоченным деле-

нием на *тему* и *рему*. *Тема* ("компонент актуального членения предложения, исходный пункт сообщения, – то, относительно чего нечто утверждается в данном предложении" [12, с.507]) преимущественно занимает начальную позицию в предложении. *Рема* ("компонент актуального членения предложения, то, что утверждается или спрашивается об исходном пункте сообщения – *теме* – и создаёт предикативность, законченное выражение мысли" [12, с.410]) занимает конечную позицию: *Первоначальный застройщик и приобретатели долей || несут по всем обязательствам... солидарную ответственность; Комитет профсоюза завода || стал больше уделять внимания проверке исполнения принимаемых решений*. Тема высказывания в подобных текстах декларативного характера преимущественно заканчивается перед первым глагольным или именным сказуемым, которое гипотетически может считаться формальным маркером начала рематического компонента.

Интересны в этом отношении позиции упомянутых выше союзов. Для союза *если*, например, характерно использование в рематической части предложения со значением аргументированности и мотивировки: *Страховая сумма выплачивается, || если в течение года ... наступит постоянная утрата трудоспособности; Ограничения перемещения товаров и услуг могут вводиться в соответствии с федеральным законом, || если это необходимо для обеспечения безопасности, защиты жизни и здоровья людей...* Союзы *в случае*, *в случае когда* используются преимущественно в составе темы предложения: *В случае несоблюдения требований, предусмотренных пунктом 1 настоящей статьи, || суд, арбитражный суд или третейский суд может отказать лицу в защите принадлежащего ему права*.

## Заключение

В 90-е годы прошлого века авторам пришлось заниматься практическим извлечением знаний из текстов для формулирования правил, которые затем были использованы в созданной нами тогда динамической экспертной системе «Сейсмо», предназначенной для построения сейсмопрогнозов [13]. Эта работа велась в рам-

ках трёхлетнего проекта сотрудничества Института физики Земли РАН, Института проблем передачи информации РАН и СААИ (Советская ассоциация искусственного интеллекта, ныне РАИИ - Российская ассоциация искусственного интеллекта). В качестве источников информации особенно большую роль сыграли научные отчеты ИФЗ РАН, в которых мы впервые столкнулись с проблематикой выделения внешней логической компоненты из текстов и преобразования полученной информации в правила, на которых основывалась работа системы «Сейсмо» [14].

В тот период нам казалось, что задача соответствующей обработки текста не является особенно трудной, если следить в тексте лишь за конструкциями типа «если ... то». Однако со временем стало ясно, что в общем случае эта задача не является простой и требует того анализа, основные направления которого мы наметили в [4] и в настоящей публикации.

## Литература

1. Мельчук И.А. Опыт теории лингвистических моделей "Смысл ↔ Текст". – М.: "Наука", 1999.
2. Виноград Т. Программа, понимающая естественный язык. – М.: "Мир", 1976.
3. Осипов Г.С. Приобретение знаний интеллектуальными системами. – М.: "Наука". Физ.-мат. лит., 1997.
4. Савинич Л.В., Стефанюк В.Л. Представление конструкций со значением обусловленности. Труды 2-й международной конференции «Системный анализ и информационные технологии (САИТ-2007)», Т. 1, с. 171-173. – М.:URSS, 2007.
5. Стефанюк В.Л., Жожикашвили А.В. Сотрудничающий компьютер: проблемы, теории, приложения. – М.: "Наука", 2007.
6. Богуславская О.Ю., Левонтина И.Б. Смыслы 'причина' и 'цель' в естественном языке. "Вопросы языкознания", № 2. – М., 2004
7. Советский энциклопедический словарь. – М.: "Советская энциклопедия", 1981.
8. Русская грамматика, т. II. Изд. АН СССР. Ин-т рус.яз. – М.: "Наука", 1980.
9. Латышева А.Н. О семантике условных, причинных и уступительных союзов в русском языке. "Вестник Моск. ун-та. Сер. 9. Филология. № 5. – М., 1982.
10. Апресян Ю.Д. Лексическая семантика. Синонимические средства языка. Изд. 2-е. – М.: "Языки русской культуры", 1995.
11. Арутюнова Н.Д. Предложение и его смысл. – М.: "Наука", 1976.
12. Лингвистический энциклопедический словарь. – М.: "Советская энциклопедия", 1990.

13. Стефанюк В.Л., Жожикашвили А.В. Создание и сопровождение базы знаний в динамической экспертной системе. Проблема "Сейсмичность". – М.: Институт физики Земли, 1992 (Препринт, 26 с.)
14. Стефанюк В.Л. Поведение квазистатической оболочки в изменяющейся нечеткой среде. Труды IV национальной конференции с международным участием "Искусственный интеллект – 94", Рыбинск, сентябрь 16-24, 1994, Т.1, с. 199-203. – Тверь: АО НИИ "Центрпрограммсистем", 1994.

**Савинич Людмила Владимировна.** Старший научный сотрудник Института проблем передачи информации им. А.А. Харкевича РАН. Окончила Московский педагогический государственный университет в 1971 году. Кандидат филологических наук. Имеет 10 публикаций. Область научных интересов: компьютерная лингвистика, лингвистика текста, прагматика.

**Стефанюк Вадим Львович.** Ведущий научный сотрудник Института проблем передачи информации им. А.А. Харкевича РАН. Окончил МГУ в 1962 году. Доктор технических наук; профессор РУДН. Имеет 192 публикации. Грамота за лучшую публикацию в журнале *Kybernetes* (Великобритания). Круг научных интересов: математическое и компьютерное моделирование мышления, искусственный интеллект, коллективное поведение систем, локально-организованные системы.