

## РОЛЬ ПУНКТУАЦИИ В РАЗРЕШЕНИИ НЕОДНОЗНАЧНОСТИ<sup>1</sup> THE ROLE OF PUNCTUATION IN DISAMBIGUATION

*Бердичевский А.С. (alexberd@yandex.ru), Московский государственный университет им. М. В. Ломоносова,  
Иомдин Б.Л. (iomdin@ruslang.ru), Институт русского языка им. В.В. Виноградова РАН*

Рассматривается роль пунктуации в разрешении неоднозначности. Знаки препинания не только организуют текст, но и передают часть заключенной в нем информации. Иногда пунктуация служит доступным и эффективным средством устранения неоднозначности. Предлагается классификация и анализ таких ситуаций.

### 1. Введение

Предмет настоящей работы – роль пунктуации в разрешении неоднозначности в тексте (фактически – в пределах предложения). Известно, что знаки препинания не только организуют письменный текст для облегчения его восприятия читателем, но и непосредственно передают часть заключенной в тексте информации. В частности, иногда пунктуация, устраняя неоднозначность, служит единственным доступным средством выбора правильной интерпретации текста.

В любом тексте скрыто гораздо больше неоднозначности, чем кажется на первый взгляд, поскольку значительную часть теоретически возможных, но практически невероятных структур человек, исходя из здравого смысла и знаний о мире и ситуации, отмечает не задумываясь. Системы автоматической обработки текста (АОТ), однако, таким здравым смыслом не обладают и потому вынуждены прибегать к иным способам устранения неоднозначности; в некоторых случаях так приходится поступать и человеку.

Насколько нам известно, большинство существующих систем АОТ мало используют пунктуацию для разрешения неоднозначности. Рассмотрим следующий пример (адаптирован из [Nunberg et al. 2002]):

(1) Mary recognised the man who entered the room, and gasped.

Запятая здесь показывает, что *gasped* относится к *Mary*, а не к *the man*. При отсутствии запятой были бы допустимы обе интерпретации. При подаче (1) система онлайн-перевода PROMT дает следующий результат:

(2) Мэри признала человека, который вошел в комнату, и задыхался.

Очевидно, что исходное предложение проинтерпретировано неправильно, т.к. не использована информация, выраженная пунктуацией. При этом нельзя сказать, что пунктуация полностью проигнорирована: по-видимому, запятая прямо перенесена в русский перевод, что приводит к пунктуационной ошибке. Если подать системе то же предложение без запятой, мы получим:

(3) Мэри признала человека, который вошел в комнату и задыхался.

На этот раз получившееся предложение не содержит ошибки и представляет собой адекватный перевод, хотя и лишь одной из двух возможных интерпретаций оригинала.

На наш взгляд, пунктуация во многих случаях представляет собой доступное средство устранения неоднозначности, заслуживающее того, чтобы им воспользоваться.

Сразу отметим, что мы хорошо представляем себе трудности, с которыми сопряжено практическое использование пунктуации в системах АОТ. Во-первых, знаки препинания сами по себе многозначны и могут сочетаться друг с другом в самых разнообразных конструкциях. Кроме того, в текстах пунктуационные ошибки встречаются, пожалуй, чаще многих других, а значит, на пунктуацию нельзя полагаться как на достаточно надежный источник. Более того, при анализе реального текста ошибочно расставленные знаки пунктуации могут привести к неверным выводам, в то время как здравый смысл и/или анализ иных структур предложения подсказывает правильный вариант.

В настоящей работе мы, отвлекаясь от этого, будем рассматривать идеальную ситуацию: пунктуация в тексте расставлена совершенно правильно. Как можно ее использовать для разрешения неоднозначности? Какие существуют теоретические возможности и как их можно применить на практике?

<sup>1</sup> Б. Л. Иомдин хотел бы выразить благодарность за финансовую поддержку Фонду содействия отечественной науке. Работа выполнена при частичной поддержке Программы фундаментальных исследований отделения историко-филологических наук РАН "Русская культура в мировой истории" и гранта Президента РФ для поддержки ведущих научных школ № НШ-5611.2006.6.

## 2. Краткий обзор литературы

В российской лингвистической литературе пунктуационная система русского языка описана довольно подробно; упомянем прежде всего [Шварцкопф 1988], а также работы А. Б. Шапиро, В. А. Ицковича, Н. С. Валгиной, Н. Н. Ореховой, сборник статей [Розина ред. 2006], справочник [Лопатин ред. 2006].

Тем не менее, нам неизвестны работы, которые были бы посвящены собственно роли пунктуации в разрешении неоднозначности, тем более в аспекте АОТ, хотя эта тема, конечно, затрагивается в вышеупомянутых работах, а также в некоторых других специальных работах, см. например [Зинькина и др. 2005] или [Pavlova et al. 1997].

В зарубежной литературе такие исследования есть, хотя, по нашим данным, они немногочисленны. В числе самых известных назовем сборник трудов, посвященных пунктуации в компьютерной лингвистике, которые были представлены на отдельном семинаре на конференции ACL/SIGPARSE в 1996 году [HCRC 1996]. Упомянем также работу [Nunberg 1990], которую многие англоязычные авторы называют основополагающей в деле исследования пунктуации, и статью [Salton 1958], одну из первых работ по использованию пунктуации в АОТ.

## 3. Юридические тексты: пунктуация и общие соображения

Существуют случаи, когда не только системы АОТ, но и человек вынужден разрешать неоднозначность, опираясь не на общие соображения, а на формальные признаки, содержащиеся в письменном тексте. Такое возможно при анализе юридического текста: договора, соглашения или закона.

Рассмотрим недавнее дело о «самой дорогой запятой в истории», конфликт канадских компаний Rogers Communications Inc. и Aliant Inc [CRTC 2006]. Между компаниями был в 2002 году заключен некоторый договор, содержавший следующий пункт:

(4) [The agreement] shall be effective from the date it is made and shall continue in force for a period of five (5) years from the date it is made, and thereafter for successive five (5) year terms, unless and until terminated by one year prior notice in writing by either party.

(5) [Договор] вступает в силу в день подписания и остается в силе в течение пяти (5) лет со дня подписания, а затем пролонгируется на пятилетние периоды, если только не будет расторгнут любой из сторон, подавшей о том письменное уведомление не позднее, чем за год до расторжения.

Тем самым, компания Rogers считала, что до 2007 года заключен нерасторжимый договор, а в 2006 году станет ясно, будет ли он автоматически продлен еще на пять лет. Однако в 2005 году от Aliant пришло письменное уведомление о расторжении договора. Конфликт возник из-за того, что Rogers считала условное придаточное *unless and until terminated...* относящимся **только** ко второй части главного предложения *and thereafter for successive five (5) year terms*, говорящей о дополнительном сроке и отделенной от первой запятой, тогда как Aliant утверждала, что условие относится ко **всему** главному предложению, т.е. как к дополнительному, так и к основному сроку. Aliant, основываясь на пунктуационной норме английского языка, аргументировала свою интерпретацию наличием в предложении **второй запятой**: если бы ее не было, то условное придаточное было бы очевидным образом тесно связано со второй частью, и тем самым не относилось бы к первой, отделенной от комплекса «вторая часть–придаточное» запятой. Rogers была вынуждена признать формальную правильность такой интерпретации, хотя в жалобе, поданной в соответствующий регулирующий орган, подчеркивала, что смысл в соглашении обеими сторонами вкладывался иной. Решение было вынесено в пользу компании Aliant, которая получила право разорвать контракт в 2006 году. Это позволило ей значительно повысить плату за свои услуги, из-за чего Rogers может потерять более двух миллионов долларов – стоимость одной запятой.

Отметим, что русский текст (5) тоже может вызвать подобные споры. Хотя для него более вероятной представляется интерпретация, при которой придаточное относится только ко второй части, нельзя полностью исключать и ту, при которой придаточное будет отнесено ко всему предложению. Лучше было бы разбить предложение на два или переформулировать его более эксплицитно.

Это один из самых скандальных, но отнюдь не исключительный пример такого рода.

Рассмотрим текст правил, устанавливающих льготы на проезд в городском транспорте в Москве:

(6) ...а также лица сопровождающие инвалидов 1-ой группы (или имеющих III степень ограничения способности к трудовой деятельности) и детей инвалидов, могут получить бесплатный посадочный талон сроком на 180 дней в киосках по продаже проездных билетов ГУП «Мосгортранс».

Приведенный отрывок недвусмысленно разрешает лицам, сопровождающим детей, чьи родители являются инвалидами, получать бесплатный посадочный талон; про состояние здоровья самих сопровождаемых детей ничего не говорится. Между тем общие соображения позволяют предположить, что на самом деле имеются в

виду лица, сопровождающие детей-инвалидов, т.е. в правилах пропущен дефис<sup>2</sup> (это подтверждается тем, что в тексте есть и другие ошибки, в том числе пунктуационная).

#### 4. Пунктуация и неоднозначность

Рассмотрим типы случаев разрешения неоднозначности с помощью пунктуации. Оговоримся сразу, что наша классификация не является исчислением всех возможных случаев, но представляет собой анализ некоторых типичных примеров.

Прежде всего выделим два основных класса: знаки препинания могут выполнять только **синтаксическую** функцию, а могут также выражать и **семантические** отношения между элементами предложения. Отметим, впрочем, что разделение даже на эти классы достаточно условно.

##### 4.1. Роль пунктуации в разрешении неоднозначности: синтаксический уровень

Роль знаков препинания на синтаксическом уровне состоит в основном в указании границ (или отсутствия границ) элементов предложения и/или их сферы действия (СД). На наш взгляд, можно выделить три часто встречающихся случая:

1. Указание правой или левой границы элемента. Сюда относится хрестоматийный случай *казнить нельзя помиловать* и подобные.

(7) Путин отстоял рождественскую службу и позвонил Ющенко.

Отсутствие запятой перед *и* показывает, что перед нами одно простое предложение, а не два: продолжаете клауза про Путина, субъект сказуемого *позвонил* – Путин, тем самым Ющенко остается только роль дополнения. Запятая ограничила бы СД подлежащего Путин справа, а сказуемого *позвонил* – слева, тогда было бы ясно, что подлежащее сказуемого *позвонил* – Ющенко.

(8) Год проработали, и нас не уволили.

Обратный случай: запятая показывает, что здесь два простых предложения, причем подлежащие в обоих опущены: скорее всего, *год проработали мы*, а *не уволили нас они*. При отсутствии запятой получилось бы, что *они* проработали (видимо, на руководящих должностях) год и за этот год нас не уволили.

По классификации Б. С. Шварцкопфа знаки пунктуации выполняют здесь **функцию разделения (отделения)** и дополняющую ее **функцию связи**.

2. Указание обеих границ элемента. Сюда относятся практически любые элементы, которые по правилам должны выделяться пунктуацией с двух сторон, например:

а) придаточные предложения, находящиеся не в конце сложного.

(9) ...встретить надежного, порядочного, умного человека с чувством юмора, на которого можно всегда положиться в трудную минуту, который знает, что хочет и умеет любить одну единственную. (Free International Dating Website).

Отсутствие запятой после *хочет* показывает, что придаточному *который знает...* следует приписать одну из двух возможных интерпретаций: *знает, что хочет любить одну-единственную и умеет любить одну-единственную* или *знает, что хочет одну-единственную и умеет любить одну-единственную*. Интерпретация *знает, что хочет; и умеет любить одну-единственную* исключается (хотя, скорее всего, на самом деле именно она имеется в виду, а автором (9) допущена ошибка). Выбрать же между двумя остающимися пунктуация не позволяет. Запятая перед *что* показывает начало придаточного, относящегося к *знает* – пока не появится следующая запятая, или другой знак, способный ее заменить (например, точка), все глаголы будут прямо или косвенно подчинены *знает*.

б) вводные слова и конструкции:

(10а) Вы верно перевели этот отрывок? [Розенталь 1996].

(10б) Вы, верно, перевели этот отрывок?

Отсутствие запятых до и после *верно* (двойной запятой) показывает, что *верно* – обстоятельство и предложение означает ‘правильно ли вы перевели этот отрывок?’. Наличие двойной запятой, как в (10б), придает предложению значение ‘правда ли, что вы перевели этот отрывок?’. Отсутствие запятых показывает более тесную связь наречия и глагола, наличие – «освобождает» наречие и позволяет ему относиться ко всему высказыванию.

(11) – Да с графом де Морсер, – отвечал Монте-Кристо. (А. Дюма, Граф Монте-Кристо, пер. с франц.).

<sup>2</sup> Мы относим включаем в число сегментных знаков препинания дефис, хотя в российской лингвистической традиции это не принято. Тем не менее, как можно видеть, в интересующем нас аспекте дефис практически совпадает по свойствам со знаками препинания.

Отсутствие запятой после *да* показывает, что это ограничительная частица, и смысл приблизительно таков: ‘с графом, с кем же еще’. Наличие запятой показало бы, что это утвердительная частица, смысл – ‘верно, с графом’.

в) обособленные приложения:

(12) Антонов, судья и Кузнецов.

Отсутствие запятой после *судья* показывает, что перечисляются три человека. Наличие показало бы, что *судья* не может являться однородным членом перечисления наряду с *Антонов* и *Кузнецов*, а значит, является обособленным приложением к *Антонов*, то есть перечислены два человека (один – судья Антонов, второй – Кузнецов). Отметим, что в английском языке пунктуация в аналогичной конструкции устроена не так:

(13) Hashimoto Ryutaro, the Prime Minister of Japan, and Bill Clinton. [Fornell 1996]

В данном случае запятая после *Japan* неинформативна, так как в английском при перечислении однородных членов постановка запятой перед последним членом, отделенным союзом *and*, допустима, но не обязательна. Таким образом, при наличии запятой могут быть перечислены как два, так и три человека, ее отсутствие однозначно бы указывало на трех. Ср. также [Kazawa et al. 2001].

По классификации Б. С. Шварцкопфа знаки пунктуации выполняют здесь **функцию выделения** и дополняющую ее **функцию связи**.

3. Более сложное структурное ограничение СД элемента.

Рассмотрим два примера (адаптировано из [Say et al. 1996]):

(14) Новое, постоянное расписание

(15) Новое постоянное расписание

В первом случае сообщается, что расписание новое и постоянное (скорее всего, старое было временным). Во втором – что имеется новое постоянное расписание (видимо, до него было другое постоянное расписание). Как именно достигается такой эффект? В (15) определения неоднородные, между ними не ставятся запятая. В СД *новое* попадает вся стоящая справа ИГ: *постоянное расписание*. В (14) запятая ограничивает СД *новое*, но «временно», она лишь сообщает, что определение *постоянное* туда не попадает, но когда перечисление однородных определений закончится (что будет отмечено отсутствием запятой!) и встретится существительное, то попадет. Ср. также [Иорданская 2000: 389–390].

Похоже был устроен и пример (4): СД условного придаточного должна быть заполнена слева. Пресловутая запятая не выводит оттуда вторую часть главного, она лишь показывает, что обе сочиненные части главного равноправны и **обе** попадают в СД придаточного.

(16) ...за ущерб, явившийся последствием обвала, оползня, оседания почвы или фундамента или эрозии, вне зависимости от их причин.

Какие формальные соображения позволяют установить, что *эрозии* не является однородным с *почвы* и *фундамента*, т.е. не подчинено *оседания*? Отсутствие запятой после *почвы* и *фундамента*, которое показывает, что уровень однородных членов кончился и мы возвращаемся на уровень их хозяина, *оседания*, с которым *эрозии* соединено сочинительной связью.

Действительно, союз позволяет убрать запятую из перечисления однородных **один** раз (если только речь не идет о простых предложениях: тогда ни разу), если же он сам начинает повторяться, то запятая ставится и перед ним. Таким образом, видя, что перед первым *или* запятой нет, мы понимаем, что однородных членов только два. Если бы запятая стояла, мы бы искали третий член (или решили бы, что соединенные союзом *или* члены суть два названия одного и того же предмета). Наличие/отсутствие запятой здесь выполняет **функцию предупреждения** по Шварцкопфу.

#### 4.2. Роль пунктуации в разрешении неоднозначности: семантико-синтаксический уровень

В этот класс попадают случаи, в которых знаки препинания не только указывают границы элементов предложения, но еще и выражают те или иные отношения между ними.

(17) – Представляешь, [такой-то] опозорился как лингвист.

– С запятой или без? (Диалог, услышанный одним из авторов).

Вопрос совершенно законен: если без запятой перед *как*, то [такой-то] опозорился в качестве лингвиста, оказался плохим лингвистом, а если с запятой, то опозорился, словно лингвист (лингвист выступает некоторым эталоном позора).

(18) Сын: Мама, это я.

Мать: Мама – это я! (Из анекдота).

Запятая показывает, что *мама* – обращение, оно должно рассматриваться вне основной структуры предложения, и *это я* можно проанализировать только одним способом. Тире же показывает, что *мама* и *я* образуют

конструкцию подлежащее–сказуемое. Ср. аналогичную ситуацию в (19):

(19) Непостоянство, имя твое – женщина! (В. Шекспир, Гамлет, пер. с англ.)

Если поменять местами запятую и тире, то *женщина* станет обращением, а *непостоянство* – частью предикации (что, кстати, будет гораздо менее необычной конструкцией, чем приведенная).

(20) Иначе все, что вам будет слышно, это «что вам угодно?» (Б. Гребенщиков, Глаз).

Кавычки здесь показывают, что *что вам угодно* является цитатой, инородным включением в текст<sup>3</sup>, и должно рассматриваться самостоятельно, вне связи с основным предложением. Если их убрать, получается: *Иначе все, что вам будет слышно, это что вам угодно*, т.е. ‘будет слышно только то, что вам угодно слышать’<sup>4</sup>.

### 5. Пунктуация в широком смысле

Для разрешения неоднозначности можно использовать не только сегментные знаки препинания.

(21) А вас, Козлов, я попрошу остаться. [Из анекдота]

Система ЭТАП-3 предлагает два перевода этого предложения на английский: 1) *And you, the Goats, I shall ask to remain*; 2) *And, Kozlov, I shall ask you to remain*. Как легко видеть, первый перевод странен, зато второй вполне адекватен. Если же ввести предложение, написав *козлов* со строчной буквы, то получим только один вариант: *And you, the goats, I shall ask to remain*. Таким образом, использование прописной буквы заставляет предположить, что лексема является именем собственным, и дает возможность определить правильную структуру предложения.

(22) Я знаю, *что* она тебе скажет. [Шварцкопф 1988]

Выделение курсивом с большой вероятностью (но не с абсолютной достоверностью) показывает, что на *что* здесь стоит фразовое ударение – тем самым это не союз, а союзное слово (смысл предложения ‘я знаю, что (именно) она тебе скажет’, а не ‘я знаю, что она тебе скажет (то-то и то-то)’).

Видно, что в принципе для разрешения неоднозначности можно использовать и пунктуацию в широком смысле: расположение текста на странице, различные средства выделения и пр. Конечно, на эти данные можно полагаться еще меньше, чем на данные собственно пунктуации: здесь нет определенных норм, и каждый текст оформляется с той или степенью произвольности. Невозможно «заложить в компьютер» строгие правила анализа типографских пунктуационных средств, т.к. нет правил их синтеза, общих для всех пишущих. Например, нельзя заранее знать, что именно курсив будет передать фразовое ударение, как в (22). Вероятно, можно использовать следующий способ: определить, какие фрагменты оформлены **не так, как** основной текст (т.е. **выделены**), а затем попытаться интерпретировать это выделение.

### 6. Использование пунктуации в системах АОТ

Мы не предлагали конкретных алгоритмов реализации пунктуационного разрешения неоднозначности для систем АОТ, т.к. они сильно зависят от общего устройства конкретной системы. В частности, мы уверены, что пунктуация может использоваться (и используется) в системах, основанных как на правилном, так и на статистическом подходах, но способы использования должны, вероятно, сильно различаться. Однако мы постарались показать, что пунктуация включает в себе ценную и – порой – не так уж сложно извлекаемую информацию, которой не стоит пренебрегать. Системы АОТ следует последовательно обучать пунктуационным правилам и соблюдению этих правил. Так, увидев отсутствие запятых в примере (16), система должна понять, что с чем сочиняется, а расстановку тире и запятых в примерах (18) и (19) система должна учитывать при установлении основы предложения.

Отдельную проблему знаки препинания представляют при переводе и особенно при машинном переводе<sup>5</sup>. Следует ли их переводить или стирать после получения структуры предложения и порождать заново по правилам выходного языка? Как можно видеть из примера (2), первая стратегия легко может привести к ошибке. С другой стороны, вторая значительно сложнее: трудности с применением пунктуационных правил хорошо знакомы носителям многих языков, можно ли надеяться полностью обучить им машину? Насколько нам известно, обычно используется сочетание обеих стратегий.

<sup>3</sup> Впрочем, это лишь одно из значений кавычек; они могут передавать и, например, метафорическое употребление, и тогда сказанное неверно. Здесь существенным показателем является также вопросительный знак: мы понимаем, что имеем дело с целым предложением.

<sup>4</sup> Отметим, что эта строка из песни часто встречается в Интернете именно в последнем варианте, т.к. на слух различить неоднозначность невозможно, для этого необходим письменный текст.

<sup>5</sup> О несоответствии пунктуации входного и выходного языков (в данном случае английского и шведского) см., например, в работе [Merkel 2001].

Иногда знаки препинания может оказаться полезно переводить лексическими средствами (так, очевидно, что русское двоеточие может часто быть дополнено *а именно* или заменено на *потому что*, запятую можно часто дополнить союзом *и* или заменить на него и т.п.). В работе [Ghazala 2005], например, отмечается, что при англо-арабском переводе знаки препинания трудно или невозможно передать пунктуационными средствами, поскольку арабская пунктуация по степени обязательности, значимости и системности не может сравниться с английской; соответственно, надо искать какие-то иные способы.

#### Список литературы

1. Зинькина Ю.В., Пяткин Н.В., Невзорова О.А. Разрешение функциональной омонимии в русском языке на основе контекстных правил // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции Диалог'2005 (Звенигород, 1–6 июня 2005 г.). М.: Наука, 2005.
2. Иорданская Л.Н. Соподчинение прилагательных в русском языке (по следам Вендлера) // Слово в тексте и в словаре. Сборник статей к 70-летию акад. Ю. Д. Апресяна. М.: Языки русской культуры, 2000.
3. Лопатин В.В. (отв. ред.) Правила русской орфографии и пунктуации. Полный академический справочник. М.: Эксмо, 2006.
4. Орехова Н.Н. Пунктуационная система языка: формирование и динамика развития (автореф. дисс. ... д.ф.н.). Воронеж, 2001.
5. Розенталь Д.Э. Справочник по орфографии и пунктуации. М., 1996.
6. Розина Р.И. (отв. ред.) Язык и мы. Мы и язык (сборник статей памяти Б. С. Шварцкопфа). М.: РГГУ, 2006.
7. Шапиро А. Б. Современный русский язык. Пунктуация. М.: УРСС, 2006.
8. Шварцкопф Б.С. Современная русская пунктуация: система и ее функционирование. М.: Наука, 1988.
9. CRTC 2006: Canadian Radio-television and Telecommunications Commission. Telecom Decision CRTC 2006-45. <http://www.crtc.gc.ca/archive/ENG/Decisions/2006/dt2006-45.htm>
10. Fornell J. Punctuation in the Bravice English-To-Japanese Machine Translation System. In: Jones B., ed. Punctuation in Computational Linguistics, UCSC, Santa Cruz, CA. SIGPARSE 1996 (Post Conference Workshop of ACL96), 1996. <http://www.cogsci.ed.ac.uk/hcrc/publications/wp-2.html>
11. Ghazala H. Stylistic-semantic and grammatical functions of punctuation in English-Arabic translation. In: Babel 50 (3): 2005: 230-245.
12. HCRC 1996: Abstracts of HCRC Workshop Proceedings. <http://www.hcrc.ed.ac.uk/publications/abstracts-wp.html>
13. Kazawa H., Isozaki H., Maeda E. NTT Question Answering System in TREC 2001. In: Voorhees E.M., Harman D.K. (eds.). NIST Special Publication 500-250. Proceedings of the Tenth Text REtrieval Conference (TREC-2001), Gaithersburg, Maryland, November 13-16, 2001. <http://trec.nist.gov/pubs/trec10/papers/nttcs-trec10-proceeding.pdf>
14. Merkel M. Comparing source and target texts in a translation corpus. In: 13th Nordic Conference on Computational Linguistics, NoDaLiDa'01, Uppsala, Sweden, 2001. <http://www.ida.liu.se/~magne/publications/merkel-comparing.pdf>
15. Nunberg G. The Linguistics of Punctuation. Stanford: Center for the Study of Language and Information, 1990.
16. Nunberg G., Briscoe T. and Huddleston R. Punctuation. In R. Huddleston, G. K. Pullam et al., The Cambridge Grammar of the English Language. Cambridge: Cambridge UP, 2002: 1723–1764.
17. Pavlova E., Pavlov Yu., Sproat R., Shih C., van Santen J. The Bell Laboratories Russian Text-to-Speech System. In: FCCL. Web Journal of Formal, Computational & Cognitive Linguistics. 1997. [http://www.kcn.ru/tat\\_en/science/fccl/pavlova.htm](http://www.kcn.ru/tat_en/science/fccl/pavlova.htm)
18. Salton G. The Use of Punctuation Patterns in Machine Translation. In: Mechanical Translation, Vol.5, No.1, July 1958.
19. Say B., Akman V. Current approaches to punctuation in computational linguistics. In: Computers and the Humanities. Vol.30, No.6, November 1996: 457–469.