# ПУНКТУАЦИОННАЯ СТРУКТУРА ХУДОЖЕСТВЕННЫХ ПРОИЗВЕДЕНИЙ И ЕЁ РОЛЬ В СИНТЕЗЕ ВЫРАЗИТЕЛЬНОЙ РЕЧИ ПО ТЕКСТУ

***Лобанов Б.М.***

(*lobanov@newman.bas-net.by*),

*Объединенный институт проблем информатики НАН Беларуси, Минск, Беларусь*

# АННОТАЦИЯ

*В работе приводятся статистические сведения о частоте употребления различных знаков препинания, а также сведения о частоте употребления простых и сложных предложений в художественных произведениях различного жанра.*

*Полученные статистические данные позволили выявить*

- *наиболее частотные знаки препинания и особенности их употребления,*
- *определить частотный и количественный состав простых и сложных предложений,*
- *формулировать особенности использования информации о знаках препинания для просодического оформления синтезированной речи по тексту.*

# Введение -1

- К настоящему времени системы синтеза речи достигли определённого совершенства и уже используются в ряде практических приложений, в том числе, при создании аудиокниг.
- Однако комфортность восприятия синтезированной речи в реальных условиях, в особенности в этом конкретном её применении, остаётся ещё далекой от удовлетворительной.
- Мировая тенденция развития речевых технологий указывает на актуальность создания систем синтеза выразительной речи (***expressive text-to-speech***).
- Исследования в области создания систем синтеза выразительной речи проводятся для ряда европейских языков: английского [1], немецкого [2], а также для синтеза речи на японском языке [3].
- Некоторые аспекты создания системы синтеза выразительной речи на русском языке были рассмотрены ранее в наших докладах [4].

# Введение -2

Понятие «выразительность речи» в лингвистике и педагогике сформировалось как междисциплинарное понятие одной из функций устной речи человека [5].

Одним из главных компонентов звуковой реализации выразительности устной речи является просодика: мелодика, ритмика и динамика речи. Эти элементы взаимодействуют, поддерживают друг друга и все вместе обуславливают её выразительность.

Пунктуационная структура текста отображается знаками препинания, которые выполняют функции

- выделения смысловых отрезков текста (абзацев, предложений и его частей, словосочетаний, слов, частей фонетического слова),
- указания на грамматические и логические отношения между словами,
- указания на коммуникативный тип предложения, его эмоциональную окраску, законченность, а также некоторые иные функции.

Знаки препинания, синтаксически оформляют текст, а при синтезе речи по тексту помогают осуществить её просодическое оформление (интонация, смысловые паузы, логические ударения).

# Введение -3

В данной работе мы коснёмся лишь одного из многих аспектов проблемы синтеза выразительной речи, а именно:

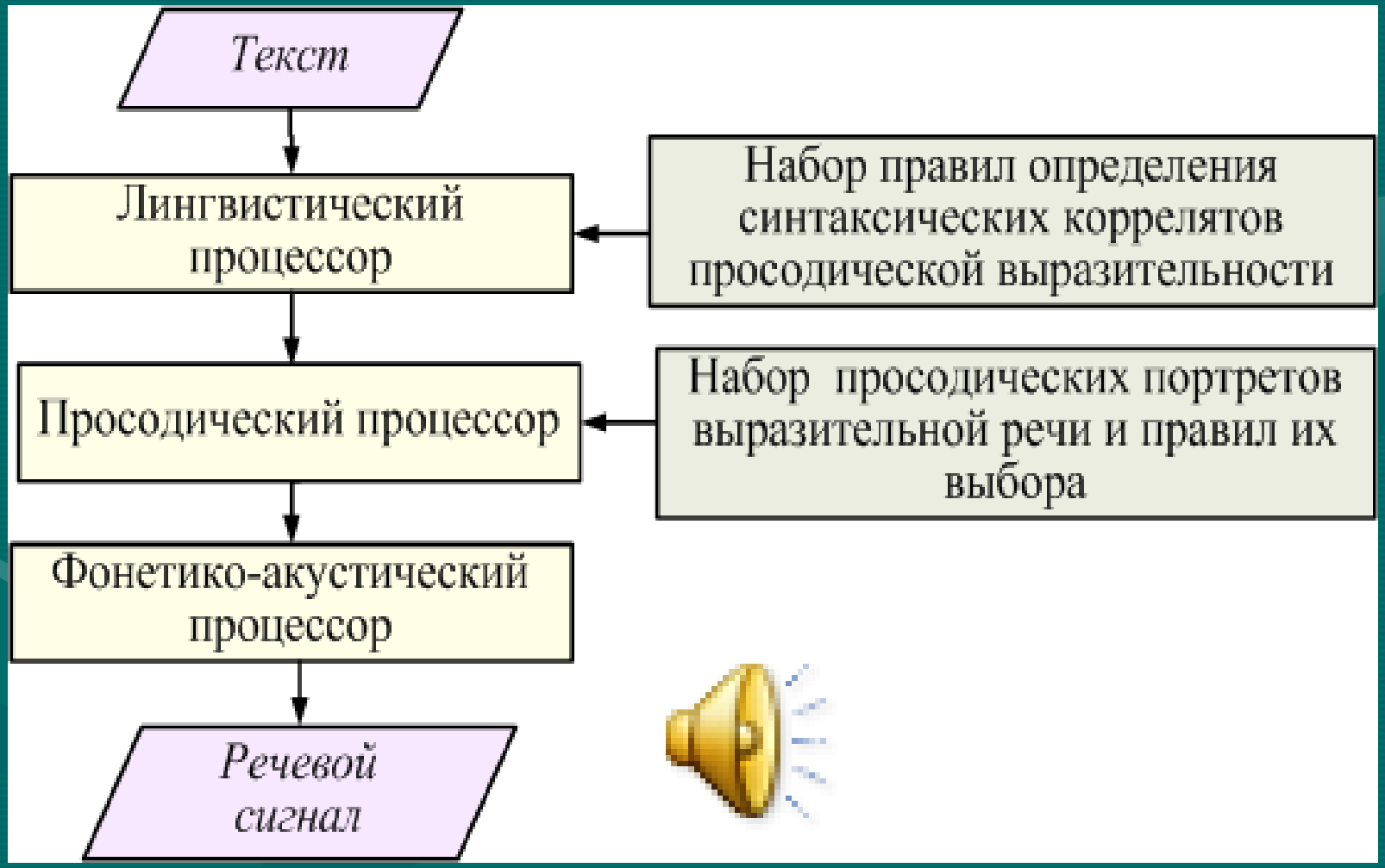
- выяснению, как воспользоваться информацией, заложенной в знаках препинания книжных текстов, для их выразительного «чтения» синтезатором речи?

С этой целью мы проанализируем пунктуационную структуру 2-х произведений художественной прозы:

- Л. Толстой “Анна Каренина”
- А.Толстой“ Пётр Первый”,
- а также полное собрание стихотворных произведений Пушкина
- и пьесу Гоголя «Ревизор».

На примере этих произведений мы рассмотрим далее статистические характеристики и роль различных знаков препинания в выражении тех или иных интонационных свойств синтезируемой речи.

# Общая структура синтезатора выразительной речи

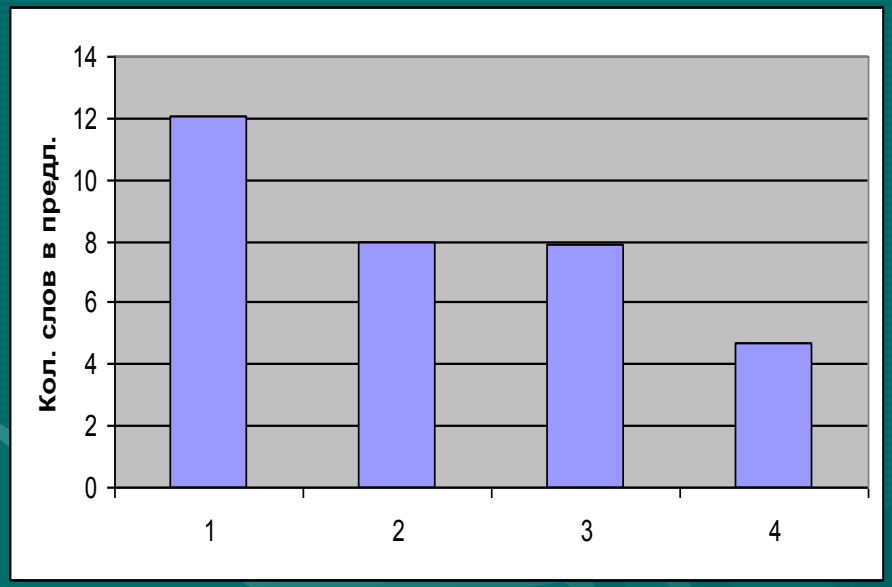


# 1. Статистика употребления знаков препинания в исследуемых текстах

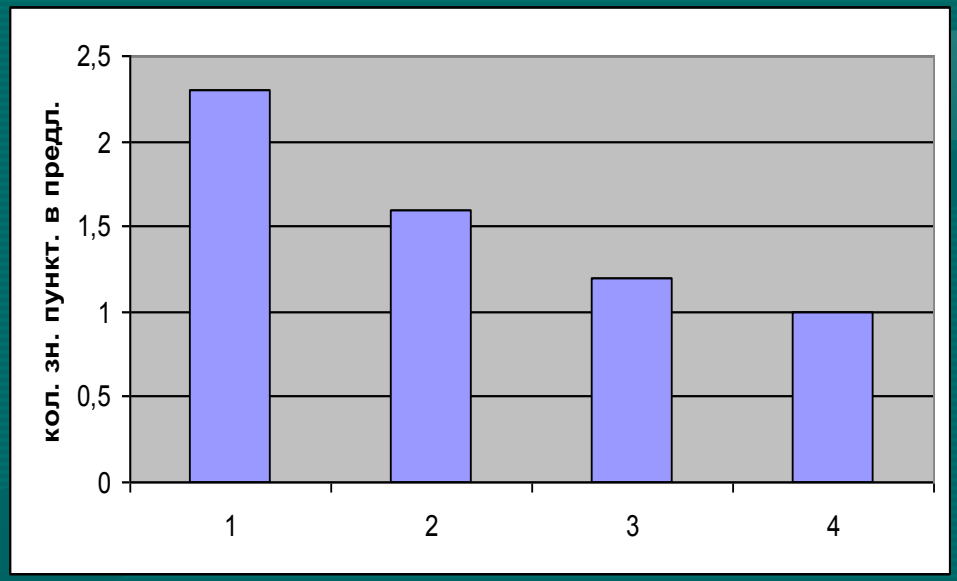
| | | Кол. слов в тексте | Кол. всех предл. в тексте | Средн кол. предл. в абзаце | Средн. кол. слов в предл. | Средн. кол. знаков путкт. в предл. | Кол. повеств. предл. в тексте (%) | Кол. воскл. предл. в тексте (%) | Кол. вопр. предл. в тексте (%) |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Л.Толстой “Анна Каренина” | 280.336 | 23.223 | 3,1 | 12,1 | 2,3 | 83,3 | 6,7 | 10 |
| 2 | А.Толстой “Пётр Первый” | 229.192 | 28.574 | 5,6 | 8,0 | 1,6 | 91,1 | 4,0 | 4,9 |
| 3 | А.Пушкин Стихи, поэмы. | 208.320 | 26.333 | 3,7 | 7,9 | 1,2 | 78,5 | 12,1 | 9,4 |
| 4 | Н. Гоголь Пьеса “Ревизор” | 20.611 | 4.365 | 3,9 | 4,7 | 1,0 | 74,6 | 15,5 | 9,9 |

# *Среднее количество слов (А) и знаков препинания (Б) внутри предложений для текстов (1 – 4)*

А)




Б)



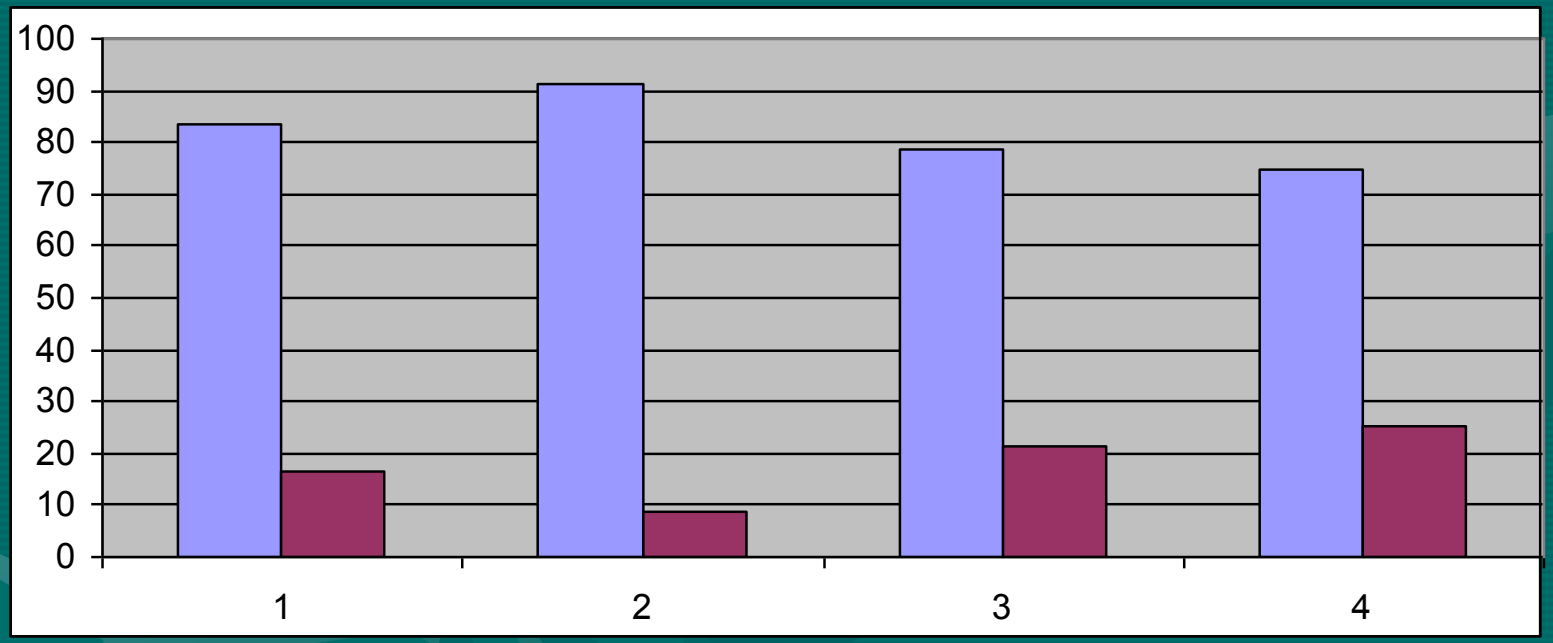

**(1) “Анна Каренина”**

**(2) “ Пётр Первый”,**

**(3) Стихи Пушкина**

**(4) Пьеса Гоголя «Ревизор».**

# *Относительное количество (в %) повествовательных (ряд 1), восклицательных и вопросительных (ряд 2) предложений в текстах 1 – 4*




**(1) “Анна Каренина”**

**(2) “ Пётр Первый”,**

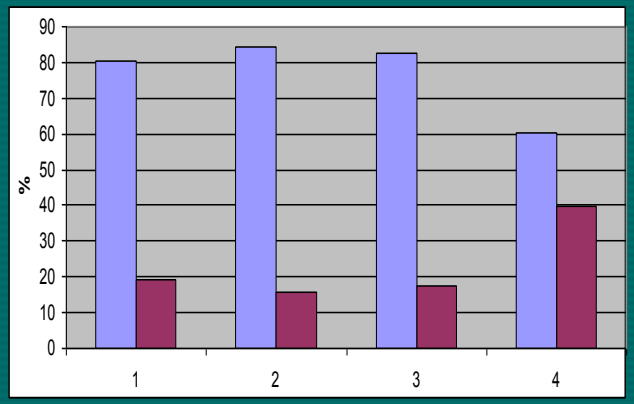
**(3) Стихи Пушкина**

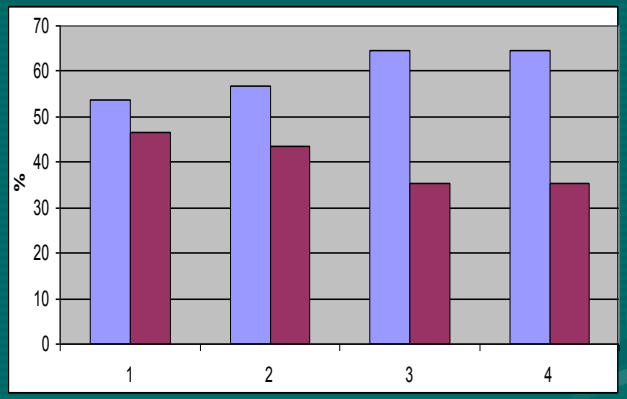
**(4) Пьеса Гоголя «Ревизор».**

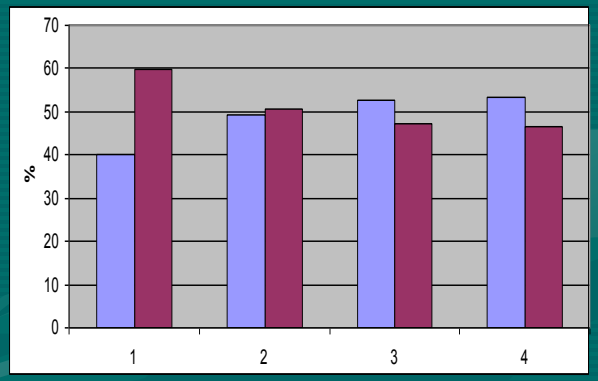
# 2. Статистика употребления простых и сложных предложений

| | | Повеств. предл. [.] | | Восклиц. предл. [!] | | Вопросит. предл. [?] | |
|---|---|---|---|---|---|---|---|
| | | сложн | простые | сложн | простые | сложн | простые |
| 1. | Анна Каренина в % | 80,6 | 19,4 | 53,6 | 46,4 | 40,1 | 59,9 |
| 2. | Пётр Первый в % | 84,3 | 15,7 | 56,6 | 43,4 | 49,4 | 50,6 |
| 3. | Стихи Пушкина в % | 82,6 | 17,4 | 64,6 | 35,4 | 52,8 | 47,2 |
| 4. | “Ревизор” Гоголя в % | 60,1 | 39,9 | 64,7 | 35,3 | 53,5 | 46,5 |

# Процентное содержание сложных (ряд 1) и простых (ряд 2) предложений в текстах 1 – 4




*А) Повествовательные* *Б) Восклицательные* *В) Вопросительные*

**(1) “Анна Каренина”**

**(2) “ Пётр Первый”,**

**(3) Стихи Пушкина**

**(4) Пьеса Гоголя**

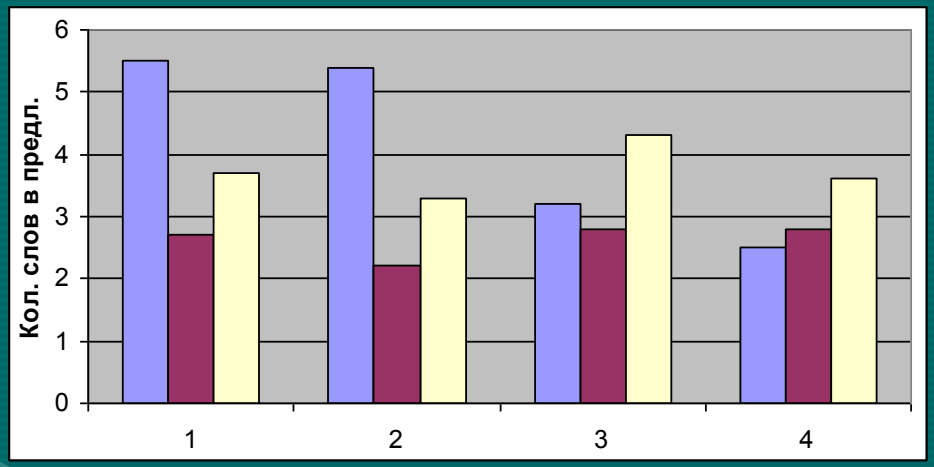
**«Ревизор».**

# Среднее и максимальное количество слов в простых предложениях

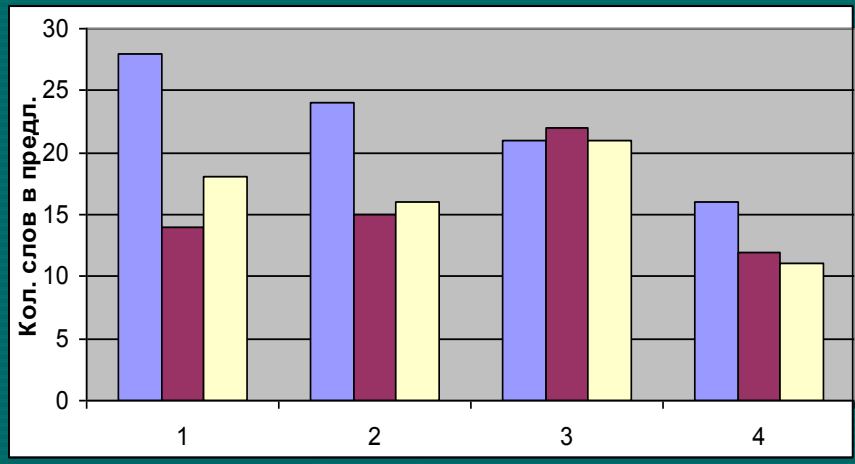
| | | Повеств. предл. [.] | | Восклиц. предл. [!] | | Вопросит. предл. [?] | |
|---|---|---|---|---|---|---|---|
| | | Средн. | Макс. | Средн. | Макс. | Средн. | Макс. |
| 1. | Анна Каренина | 5,5 | 28 | 2,7 | 14 | 3,7 | 18 |
| 2. | Пётр Первый | 5,4 | 24 | 2,2 | 15 | 3,3 | 16 |
| 3. | Стихи Пушкина | 3,2 | 21 | 2,8 | 22 | 4,3 | 21 |
| 4. | Ревизор “Гоголя” | 2,5 | 16 | 2,8 | 12 | 3,6 | 11 |

# *Среднее (А) и максимальное (Б) количество слов в простых повествовательных (ряд 1), восклицательных (ряд 2) и вопросительных (ряд 3) предложениях для 4-х текстов*

А)




Б)




**(1) “Анна Каренина”**

**(2) “ Пётр Первый”,**

**(3) Стихи Пушкина**

**(4) Пьеса Гоголя «Ревизор».**

# Примеры простых и сложных предложений максимальной длинны из романа “Анна Каренина”

| Тип предложения | Примеры |
|---|---|
| **[.] простые** | И татарин с развевающимися фалдами над широким тазом побежал и через пять минут влетел с блюдом открытых на перламутровых раковинах устриц и с бутылкой между пальцами. |
| **[.] сложные** | Ему казалось, что при нормальном развитии богатства в государстве все эти явления наступают, только когда на земледелие положен уже значительный труд, когда оно стало в правильные, по крайней мере в определенные условия; что богатство страны должно расти равномерно и в особенности так, чтобы другие отрасли богатства не опережали земледелия; что сообразно с известным состоянием земледелия должны быть соответствующие ему и пути сообщения, и что при нашем неправильном пользовании землей железные дороги, вызванные не экономическою, но политическою необходимостью, были преждевременны и, вместо содействия земледелию, которого ожидали от них, опередив земледелие и вызвав развитие промышленности и кредита, остановили его, и что потому, так же как одностороннее и преждевременное развитие органа в животном помешало бы его общему развитию, так для общего развития богатства в России кредит, пути сообщения, усиление фабричной деятельности, несомненно необходимые в Европе, где они своевременны, у нас только сделали вред, отстранив главный очередной вопрос устройства земледелия. |
| **[!] простые** | И как я смел соединять мысль о чем-нибудь не невинном с этим трогательным существом! |
| **[!] сложные** | Так же буду сердиться на Ивана кучера, так же буду спорить, буду некстати высказывать свои мысли, так же будет стена между святая святых моей души и другими, даже женой моей, так же буду обвинять ее за свой страх и раскаиваться в этом, так же буду не понимать разумом, зачем я молюсь, и буду молиться, - но жизнь моя теперь, вся моя жизнь, независимо от всего, что может случиться со мной, каждая минута ее - не только не бессмысленна, как была прежде, но имеет несомненный смысл добра, который я властен вложить в нее! |
| **[?] простые** | Неужели эта трогательная радость его при ее приближении была причиной охлаждения Анны Павловны? |
| **[?] сложные** | Разве не молодость было то чувство, которое он испытывал теперь, когда, выйдя с другой стороны опять на край леса, он увидел на ярком свете косых лучей солнца грациозную фигуру Вареньки, в желтом платье и с корзинкой, шедшей легким шагом мимо ствола старой березы, и когда это впечатление вида Вареньки слилось в одно с поразившим его своею красотой видом облитого косыми лучами желтеющего овсяного поля и за полем далекого старого леса, испещренного желтизною, тающего в синей дали? |

# *Примеры простых и сложных предложений максимальной длинны из стихотворений Пушкина*

| Тип предлож. | Примеры |
|---|---|
| **[.] простые** | В теченьи медленном река Вблизи плетень из тростника Волною сонной омывала И вкруг него едва журчала При легком шуме ветерка. |
| **[.] сложные** | Цветы, луга, ручей живой, Счастливый грот, прохладны тени, Приют любви, забав и лени, Где с Анжеликой молодой, С прелестной дщерью Галафрона, Любимой многими - порой Я знал утехи Купидона. |
| **[!] простые** | Не дай мне бог сойтись на бале Иль при разъезде на крыльце С семинаристом в желтой шале Иль с академиком в чепце! |
| **[!] сложные** | А ты, повеса из повес, На шалости рожденный, Удалый хват, головорез, Приятель задушевный, Бутылки, рюмки разобьем За здравие Платова, В козачью шапку пунш нальем - И пить давайте снова! |
| **[?] Простые** | Когда в сияньи возгорится Светильник тусклый юных дней И мрачный путь мой озарится Улыбкой спутницы моей? |
| **[?] сложные** | Давно ли с трепетом народы Несли мне робко дань свободы, Знамена чести преклоня; Дымились громы вкруг меня, И слава в блеске над главою Неслась, прикрыв меня крылом? |

# 3. Статистика употребления знаков препинания внутри и на концах предложений

| *Среднее количество знаков препинания внутри предложений* | [ , ] | [ – ] | [ , – ] | [ ; ] | [ : ] | [ ( ] |
|---|---|---|---|---|---|---|
| “Анна Каренина” Л.Н. Толстого | **1,69** | **0,43** | **0,12** | **0,05** | **0,03** | **0,02** |
| “Пётр Первый”А.Н. Толстого | **1,01** | **0,37** | **0,1** | **0,07** | **0,08** | **0,04** |
| Стихи, поэмы А.С. Пушкина | **0,78** | **0,06** | **0,01** | **0,14** | **0,1** | **0,06** |
| Пьеса Н.В. Гоголя “Ревизор” | **0,64** | **0,04** | **0,05** | **0,03** | **0,06** | **0,14** |

# Процентные соотношения предложений с различными знаками препинания в конце предложения

| | [ . ] | [ \|\| ] | [ ! ] | […] | [?] | [ \|\| –] | [ . – ] | [? – ] | […. –] | [ ! – ] | [ !.. ] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Текст 1** | **74** | **12** | **5,4** | **3,4** | **7,3** | **20** | **4,9** | **2,7** | **0,4** | **1,3** | **0** |
| **Текст 2** | **81** | **8** | **5,4** | **8,7** | **4,5** | **10** | **1,3** | **0,4** | **0,1** | **0,6** | **0** |
| **Текст 3** | **68** | **26** | **11,8** | **10** | **9,2** | **0,6** | **0,3** | **0,2** | **0,1** | **0,3** | **0** |
| **Текст 4** | **60** | **26** | **15,2** | **13** | **9,3** | **0** | **0,1** | **0,6** | **2,4** | **0,3** | **1,2** |
| **Средн.** | **70,7** | **18,0** | **9,5** | **8,8** | **7,6** | **7,5** | **1,6** | **1,0** | **0,7** | **0,6** | **0,3** |

# 4. Использование информации о знаках препинания при синтезе выразительной речи по тексту

В порядке уменьшения относительной длительности межсинтагменной паузы знаки препинания можно условно разбить на следующие группы:

- (2 – 5 сек) – суперабзац, отделяющий отдельные части произведения (параграф, глава и др.) – **[ ||| ]**.
- (1 – 2 сек) – абзацы, соответствующие разделам повествовательной – **[ || ]** и диалоговой речи **[ || –]**.
- (0,5 – 1 сек) – повествовательные – **[ . ], [ . – ], [ …], […. – ],** восклицательные – **[ !], [ ! …], [ !!!]** и вопросительные предложения – **[ ? ], [? – ], [?!]**.
- (0,2 – 0,5 сек) – вводные предложения – **[ ( ) ], [ – / – ],** а также части предложения, отделяемые знаками – **[ ; ], [ : ], [ – ]**.
- (0 – 0,2 сек) - части предложения, отделяемые запятыми.

Примером нулевой длительности паузы является запятая перед вводным словом: « *И, наконец, он сказал*».

Знаки препинания на концах предложений:

- **[ | | |], [ | | ], [ . ], [ . – ], [ ) ], […], […–]**

  **–** определяют тот или иной подтип интонации завершенности,

знаки:

- **[ !!! ], [ ! ], [ ! – ], [ !.. ]**

  **–** тот или иной подтип восклицательной интонации,

а знаки:

- **[?!], [?], [? – ]** – вопросительной.

Знаки препинания внутри предложений:

- **[ ; ], [ : ], [( ], [ , ], [ – ], [, – ]**

  **–** определяют тот или иной подтип интонации незавершенности.

Особенно разнообразную роль в реализации интонации незавершенности играют знаки **[ , ]** и **[ – ]**.

# Заключение

- Информация о знаках препинания в тексте является безусловно полезной, но явно недостаточной для синтеза выразительной русской речи.
- Дальнейшее улучшение выразительности синтезированной речи может быть достигнуто за счёт качественного членения на просодические синтагмы и расстановки акцентов внутри синтагм.
- Очевидным путём улучшения качества членения на просодические синтагмы и расстановки акцентов внутри синтагм является использование механизмов глубокого синтаксического анализа предложений, что и явится предметом наших дальнейших исследований.

# СКАЗКИ Пушкина и Ершова

- О царе Салтане
- О золотой рыбке
- О золотом петушке
- О мёртвой царевне
- О попе и балде
- О Коньке-Горбунке

**Запятая** может играть различную функциональную роль в предложении:

- отделение однородных членов предложения;
- отделение предикативных частей бессоюзного сложного или сложносочиненного предложения;
- отделение главного и придаточного предложений;
- при выделении обособленных членов предложения (причастий, деепричастий).

В зависимости от этой роли реализуется тот или иной подтип интонации незавершенности.

**Тире** в предложении также употребляется в разнообразной функциональной роли:

- отделение подлежащего и сказуемого в эллиптических предложениях;
- отделение однородных членов предложения перед обобщающим словом;
- для выражения неожиданности, резкой противоположности между членами предложения;
- для выделения вводных и вставных предложений;
- для выделения слов автора при прямой речи.