

# Автоматическая расстановка пауз в системе синтеза русской речи по тексту

Ольга Хомицевич ([khomitsevich@speechpro.com](mailto:khomitsevich@speechpro.com))

Михаил Соломенник ([solomennik-m@speechpro.com](mailto:solomennik-m@speechpro.com))

ООО «Центр речевых технологий», Санкт-Петербург

*Диалог 2010, 26-30 мая 2010 г.*

# Синтез речи по тексту

- Нормализация текста (приведение всех обозначений к обычным словам).
- Расстановка ударений, снятие омографии.
- **Деление на синтагмы, расстановка пауз.**
- Определение интонационных контуров.
- Транскрипция.
- Озвучивание.

# Зачем нужен алгоритм деления на синтагмы?

- Большие отрезки текста без пауз плохо воспринимаются на слух.
- Знаки препинания не всегда дают достаточно информации для расстановки всех пауз:
  - В тексте встречаются большие отрезки без знаков препинания.
  - Не всегда знак препинания указывает на паузу.

# Расстановка пауз в системе синтеза речи “VitalVoice”

- Развитие базового алгоритма расстановки пауз (система синтеза речи «Оратор»).
- Частичный синтаксический анализ предложения.
- Выявление неделимых групп слов, прогнозирование разрывов синтагм между ними.

# Принцип работы алгоритма

- Ищем группы связанных между собой слов.
- Частеречные «шаблоны» - заранее заданные последовательности частей речи и/или грамматических форм, соответствующие часто встречающимся в текстах группам слов.
- Шаблоны сопоставляются последовательностям слов в предложении.
- Между этими словами прогнозируется наличие синтаксической связи.
- Для постановки пауз ищутся места с наиболее слабой связью между словами.

# Пример шаблона

- V + I + PREP(accord=4) + N
- *Представитель администрации президента предпочел воздержаться от комментариев на эту тему.*
  - В базовой версии алгоритма учитывалось количество пересекающихся шаблонов для каждой пары слов, а также длины шаблонов.
  - При пересечении нескольких длинных шаблонов пауза не ставилась.
  - Возникало много ошибок, когда длинные отрезки текста не делились на синтагмы.

# Оптимизация алгоритма

- Учет разных типов связей:
  - Связь первого типа - может быть разорвана, если соответствующая позиция между двумя словами будет определена как возможное место паузы (подлежащее + сказуемое, существительное + предлог и т.п.);
  - Связь второго типа - не может быть разорвана (предлог + существительное, согласованное определение + существительное и т.п.).
  - Примеры:
    - N(case=acc) + V (*особо важную роль играет своевременное выполнение*)
    - FULL\_ADJ(accord=2) & N (*особо важную роль играет своевременное выполнение*)
- При выборе места паузы учитывается:
  - Количество шаблонов, в которые входят соседние слова.
  - Тип связи в шаблоне, связывающем слова.

# Оптимизация алгоритма

- Оптимизация набора шаблонов:
  - Уменьшено количество шаблонов, которые часто разрываются паузой.
    - Например, группы, содержащие:
      - подлежащее и сказуемое  
*Депутат областного парламента Сергей Иванов / перешел на работу к новому губернатору.*
      - предложные группы  
*Дизайнер продемонстрировал солидарность / с британскими и американскими модельерами.*
  - Увеличено количество шаблонов, отражающих дистанционные синтаксические отношения.
    - Например:
      - Поврежденная при двух взрывах шахта будет полностью восстановлена; Чувствительная к детскому горю ведущая предложила позвонить осиротевшему мальчику.



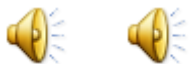
# Оптимизация алгоритма

- Учет специальных случаев.
  - Однородные члены предложения (например, «пять, шесть или семь»);
  - Вводные слова («например», «наверно», «в частности» и т.п.);
  - Сложные предлоги, энклитики, послелогои, союзы («несмотря на», «спустя», «уж», «бы» и т.п.);
  - Обозначения дат, времени (например, «двадцатого мая две тысячи десятого года»);
  - Неразрывные идиоматические последовательности («задом наперед», «бог весть» и т.п.)
  - И т.п.




# Пример работы алгоритма

По & сравнению & с & предыдущими & семью & неделями	рост + несколько & замедлился.	
по & сравнению		PREP(accord=2) & N
	предыдущими & семью & неделями	FULL_ADJ(gender=3, accord=3) & NUM(accord=3) + N
	предыдущими & семью &	FULL_ADJ(gender=2, accord=2) & NUM
	& семью & неделями	NUM(accord=2) & N
		N(case=nom, gender=2, number=2) + V(person=thrd)
		+ несколько & замедлился ADV & V

- Найдено 6 групп.
- «По сравнению с» обрабатывается как сложный предлог.
- Пауза ставится там, где нет никаких связей.



# Примеры чтения текстов

- 23 депутата Госдумы проигнорировали Медведева / и отказались подавать декларации о доходах. // В соответствии с указом президента Дмитрия Медведева / сведения об имуществе и доходах депутатов Госдумы / должны были быть опубликованы до 14 мая. // Однако декларации 23 парламентариев / на официальном сайте нижней палаты так и не появились. 
- Крупнейшие инвестиционные банки мира / приняли добровольное решение / о раскрытии части деловой информации. // Согласно данным Ассоциации финансовых рынков Европы, / речь идет о торговых операциях компаний, / которые были совершены в так называемых "темных омутах" - / закрытых торговых системах, / позволяющих банкам / обмениваться крупными пакетами ценных бумаг / без ведома регулятора / и остальных участников рынка. 
- Скандально известная американская киноактриса Линдси Лохан / наказана за задержку на Каннском кинофестивале / и неявку в срок в суд Лос-Анджелеса. // Лохан обязана еженедельно сдавать анализы / на содержание алкоголя и наркотиков / и носить специальный браслет, / фиксирующий содержание спиртного в организме. 

# Результаты: параметры оценки

- Правильность расстановки пауз в тексте сложно оценить статистически:
  - Один и тот же текст часто не делится на синтагмы однозначно;
  - Дикторская вариативность.
- Способ оценки правильности деления:
  - Правильность разрывов (BC=breaks correct) – отношение правильно определенных мест разрывов к общему количеству разрывов в тексте;
  - Правильность связей (JC=junctures correct) – правильность определения наличия/отсутствия разрыва для каждой пары слов;
  - Вставка разрыва (JI=juncture insertion) – «лишние» разрывы, вставленные алгоритмом;
  - Удаление разрыва (JD=juncture deletion) – разрывы, не найденные алгоритмом.

# Результаты: корпус спонтанной речи

Связи, %	Разрывы, %	Ошибки – вставка, %	Ошибки – удаление, %
<b>Система «Оратор»</b>			
92,05	78,12	4,08	3,87
<b>Система «VitalVoice»</b>			
93,19	83,48	2,33	4,48

- Сравнение с данными, опубликованными в [5].
- Количество ошибочно поставленных пауз сократилось почти в два раза, при незначительном росте количества мест, оставленных без пауз.

# Результаты: чтение новостных текстов

	Связи, %	Разрывы, %	Ошибки – вставка, %	Ошибки – удаление, %
<b>Система «Оратор»</b>				
Диктор 1	86,58	65,14	5,39	8,03
Диктор 2	87,96	69,60	5,56	6,48
<b>Система «VitalVoice»</b>				
Диктор 1	89,59	70,03	3,47	6,94
Диктор 2	90,59	73,98	3,83	5,57

- В новой системе паузации:
  - количество ошибочно поставленных пауз сократилось в полтора раза;
  - в среднем на 15% сократилось количество мест, ошибочно оставленных без пауз.

# Параметры оценки с учетом обязательных / необязательных пауз

- Исследования одних и тех же текстов, прочитанных разными дикторами, показывают, что во многих позициях наличие/отсутствие паузы не является обязательным.
- Модифицированная система оценки – три вида позиций в тексте:
  - обязательная пауза,
  - обязательное место отсутствия паузы,
  - возможное место паузы.
- Если пауза, сгенерированная системой, попадала в место возможной паузы, то эта пауза считалась правильной, также правильной считалась ситуация, если в месте возможной паузы пауза отсутствовала.
- Оценка: параметр ВС не рассчитывается, а вместо него вводятся два дополнительных параметра: СР - количество правильно поставленных пауз от общего количества проставленных пауз и СЈ - количество правильных мест без пауз от общего количества мест, оставленных без пауз.

## Результаты с учетом необязательных мест пауз (художественный текст)

Связи, %	Правильные – наличие, %	Правильные – отсутствие, %	Ошибки – вставка, %	Ошибки – удаление, %
<b>Система «Оратор»</b>				
97,88	98,84	97,57	1,84	0,28
<b>Система «VitalVoice»</b>				
98,90	99,59	98,68	1,00	0,10

- Количество ошибочно поставленных пауз в новой системе паузации сократилось почти в два раза.
- Почти в три раза сократилось количество мест, ошибочно оставленных без пауз.



# Проблемы и задачи для дальнейшей работы

- Проблемы частичного синтаксического анализа:
  - некоторые дистанционные синтаксические отношения определить невозможно;
  - не учитывается взаимное расположение выделенных синтаксических групп.
- Синтаксическая неоднозначность:
  - Пример: предложные группы.  
*Люксембург поддерживает кандидатуру России /(?)/ для вступления /(?)/ во Всемирную торговую организацию*

## Выводы и заключение

- Разработанный алгоритм расстановки пауз:
  - Дает хорошие результаты при обработке текстов различного типа;
  - Легко модифицировать и отлаживать;
  - Не затратный по вычислительным ресурсам.
- На настоящий момент существенно улучшены результаты работы алгоритма по сравнению с базовой версией.

# Библиография

1. Atterer M. Assigning Prosodic Structure for Speech Synthesis: A Rule-based Approach // Proc. of Prosody 2002, Aix-en-Provence. pp. 147-150.
2. Black, A.W., Taylor, P. Assigning phrase breaks from part-of-speech sequences // Computer Speech & Language, Volume 12, Number 2, April 1998. New York: Academic Press - pp. 99-117.
3. Bondarko, L.V. [et al.] Phonetic Properties of Russian Spontaneous Speech // Proceedings of the 15th ICPHS. Barcelona, Spain - 2003. pp. 2973–2976.
4. Oparin, I. Flexible rule-based breaks assignment for Russian // SPECOM 2005 proceedings: 10th international conference Speech and computer : 17-19 October, 2005. Patras, Greece. - Moscow: Moscow State Linguistics University, 2005. - pp. 293-296.
5. Oparin, I. Robust rule-based method for automatic break assignment in Russian texts // Text, speech and dialogue: 8th international conference, TSD 2005: Karlovy Vary, Czech Republic, September 12-15, 2005 : proceedings. - Berlin: Springer, 2005. pp. 356-363.
6. Вольская Н., и др. Синтезатор русской речи по тексту нового поколения // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции "Диалог'2005" (Звенигород, 1-6 июня, 2005 г.). Под ред. И.М. Кобозевой, А.С. Нариньяни, В.П. Селегея. М.: Наука, 2005. С. 234-237.
7. Кривнова О.Ф., Чардин И.С. Паузирование в естественной и синтезированной речи. – Донецк, 2002. - <http://www.russian.slavica.org/article9348.html>.
8. Опарин И.В. Автоматическое интонационное членение для решения прикладных задач // Интегральное моделирование звуковой формы естественных языков. СПб., 2005. С. 54-64.
9. Система синтеза русской речи по тексту VitalVoice: <http://vitalvoice.speechpro.com/>



**Спасибо за внимание!**