

О. Ф. Кривнова, О. С. Смирнова
г. Москва, МГУ имени М. В. Ломоносова

ТЕКСТОВЫЕ ХАРАКТЕРИСТИКИ СЛОВОРАЗДЕЛА В УСТНОЙ РЕЧИ: БАЗА ДАННЫХ КАК ИНСТРУМЕНТ ПРОГНОЗИРОВАНИЯ ПРОСОДИЧЕСКОГО ЧЛЕНЕНИЯ ЗВУЧАЩЕГО ТЕКСТА¹

Цель работы, о которой идет речь в настоящем докладе, заключалась в исследовании просодического членения (ПЧ) русской звучащей речи с целью получения новых и статистически надежных данных о природе этого явления, фонетической реализации и контролирующих факторах: коммуникативно-синтаксических, фонетических и физиологических. Важным моментом теоретической установки исследования является признание слова основной рабочей единицей как при порождении, так и при восприятии любого текста, и письменного, и устного. С точки зрения возможностей ПЧ любая граница между словами (словораздел) в тексте имеет определенный сегментирующий *потенциал*, который может реализоваться с разной вероятностью и с разной силой в зависимости от типа дискурса, контекстных условий и различных контролирующих факторов локальной и глобальной природы. Носитель языка, даже не имеющий специального лингвистического образования, согласится с тем, что соседние слова в тексте в разной степени связаны между собой по смыслу, синтаксически и даже фонетически. Этот факт можно интерпретировать как признание разной сегментирующей силы (глубины) словоразделов. В *письменном* тексте в качестве формальных показателей сегментирующей силы словоразделов (wordbreak strength) выступают знаки препинания: их наличие/отсутствие и тип. Знаки препинания не только членят текст на когерентные фрагменты, но и указывают в определенной степени на их иерархический статус. В *устной* речи аналогичную функцию выполняют просодические средства: паузы, переломы тона и другие фонетические явления на граничных участках соседних слов в последовательности. Просодически маркированные словоразделы образуют просодические швы (разрывы) в звучащем тексте, что хорошо отражает англоязычный термин «prosodic break». Естественно предположить, что внутренняя иерархия ПЧ на фразовом уровне находит отражение в разной глубине («силе») просодических швов (далее ПШ), которая создается использованием разных просодических средств между и на краях просодических составляющих. В интонационной фонологии многие исследователи разделяют точку зрения, согласно которой иерархический статус просодической составляющей однозначно соответствует глубине ПШ, завершающего эту составляющую. Это положение т.н. строгой поуровневой гипотезы (Strict Layered Hypothesis SLH) разделяется, однако, не всеми интонологами, но, к сожалению, никогда не проверялось экспериментально на сколько-нибудь представительном речевом

¹ Исследование проведено при поддержке гранта РФФИ 19-012-00534 А.

материале, см. об этом [1; 2; 2; 3]. В то же время нельзя не отметить, что количественная оценка сегментирующей силы словоразделов, или, что то же самое, глубины ПШ в озвученном речевом высказывании, создает основу для реконструкции иерархической структуры его просодических составляющих.

В отечественной лингвистике впервые обратил внимание на ПЧ и его особую природу академик Л. В. Щерба. Он писал: «В европейских языках (а вероятно и во многих других) самым могучим средством выражения связи между словами и группами слов является “интонация”, “фразировка” в самом широком смысле слова». Щерба обозначил практически все отличительные особенности этого явления, которые в настоящее время являются объектом исследования во многих работах по фразовой просодии, однако не описаны и не объяснены полностью ни для одного из европейских языков. В своих работах он использовал 4 маркера для фразовых ПШ разной глубины. В приводимых там примерах использовалась пятибалльная количественная шкала: 0, 1, 2, 3, 4. Много текстовых примеров с интроспективной разметкой ПШ содержится и в книге Р.И. Аванесова «Русское литературное произношение» [4]. В его примерах используется пять маркеров глубины ПШ, т.е. фактически шестибалльная количественная шкала.

Результаты перцептивных экспериментов по оценке сегментирующей силы словоразделов на материале разных языков (английского, нидерландского и русского) свидетельствуют о поведенческой устойчивости оценок, полученных с использованием количественной шкалы, различающей не более 5 уровней глубины ПЧ. Видимо, данная шкала отражает какие-то универсальные свойства перцептивных ощущений человека в анализируемом фонетическом пространстве. Для многих практических задач достаточно учитывать этот результат и даже сознательно использовать более крупную трехуровневую шкалу, как это советует делать А. Сандерман [2] на начальной стадии технологических разработок в области синтеза и распознавания речи. Однако для подтверждения указанной гипотезы с общefonетических научных позиций необходимо увеличить как количество и разнообразие рассматриваемых языков, так и число испытуемых, привлекаемых к ее верификации.

Материал и методика исследования. Учитывая многообразие форм устного дискурса, многослойность и многоаспектность проблемы ПЧ, речевой материал нашего проекта был ограничен русскими прозаическими текстами, озвученными в режиме чтения, с некоторыми дополнениями в виде текстов, специально озвученных для целевых фонетических экспериментов. Основной массив корпуса включает два текста: 1. фрагмент из повести И. Грековой "Кафедра" (2700 словоупотреблений = словоразделов в чтении 2-х непрофессиональных дикторов, далее текст ИГ) 2. прозаические тексты из книги Р. И. Аванесова "Русское литературное произношение" (2500 словоупотреблений = словоразделов в чтении 3-х непрофессиональных дикторов, далее текст РИА). Во всех случаях запись озвученных производилась на качественной цифровой аппаратуре в студийных условиях.

Тексты основного массива были выбраны для нашего корпуса не случайно: текст ИГ входит в набор текстов, имеющих синтаксическую разметку в формате анализатора ЭТАП-3, а тексты РИА содержат авторскую разметку просодического членения. В то же время при включении этих текстов в состав корпуса потребовалась определенная унификация лингво-просодических данных.

Фонетическая составляющая звукового корпуса включает перцептивную разметку, в которой каждому словоразделу поставлен в соответствие количественный субъективный показатель его сегментирующей силы, соответственно глубины ПШ. Разметка осуществлялась с использованием фиксированной 5-тибалльной шкалы, согласованной с результатами перцептивных экспериментов, проведенных с привлечением фонетистов-экспертов и обычных носителей языка в качестве аудиторов [5, 6]. Кроме этого, была также произведена и верифицирована полная паузальная разметка звуковых записей корпуса с одновременным измерением следующих временных показателей: длительности паузы на словоразделе, длительности предпаузального и постпаузального речевых фрагментов.

Кроме просодической разметки, в состав фонетических признаков словораздела входят также акустические характеристики (длительность, интенсивность, ЧОТ, спектральные показатели коартикуляции) звуковых сегментов в непосредственной окрестности каждого словораздела. Для получения этих характеристик была произведена ручная сегментация пограничных областей для каждого словораздела и автоматический акустический анализ выделенных звуковых сегментов.

Синтаксическая составляющая корпуса формировалась постепенно, и вопрос о выборе синтаксического формализма, наиболее адекватного для нашей задачи, до сих пор требует доработки. Текст ИГ, исходно размеченный помощью анализатора ЭТАП-3 в формате деревьев зависимостей для каждого предложения текста, был программно переведен в скобочную форму синтаксического представления предложения в терминах непосредственных составляющих. Скобочная форма более удобна для представления синтаксической информации в табличной форме и позволяет вычислять и использовать простые и осмысленные для синтаксической интерпретации ПЧ признаки, как то: плотность (сгущение) синтаксических скобок на словоразделе, соотношение закрывающих и открывающих скобок, тип объемлющей составляющей по обе стороны от словораздела и т.п. Кроме того, для ИГ были произведены две альтернативные ручные разметки синтаксиса. В одной из них использовался метод Гаспарова–Скулачевой [7], который применяется для исследования глубины синтаксических границ в структуре стихотворного текста. Этот тип разметки основан на классической синтаксической теории членов предложения и эмпирически установленной иерархии силы синтаксической связи между словами, разделенными словоразделом. Второй тип ручной разметки можно условно назвать синтактико-пунктуационным: для получения данных о влиянии типа синтаксической составляющей на ПЧ

было выделено 12 типов потенциально релевантных синтаксических структур, а затем была проведена ручная разметка текста маркерами таких составляющих. Во многих случаях границы таких составляющих в тексте маркируются в соответствующих словоразделах знаками препинания, согласно действующим в русском языке пунктуационным правилам¹, однако далеко не всегда. Таким образом, синтаксический фактор влияет на паузы не только опосредованно через знаки препинания, но должен учитываться и независимо от них. Тексты РИА были размечены вручную с выделением наиболее значимых и крупных синтаксических границ: границы текста, конец абзаца внутри текста, конец самостоятельного предложения внутри абзаца, конец элементарной клаузы внутри самостоятельного предложения, границы пояснительных и сравнительных оборотов внутри элементарной клаузы. Значимость этой разметки для ПЧ также исследовалась статистически.

На основе комплексного синтактико-просодического анализа собранного материала был разработан макет базы данных для прогнозирования места и глубины просодического членения на словоразделах звучащего текста (далее БДС). В ее структуре имеются четыре зоны: текстовая зона, зона левого контекста словораздела, центральная зона собственно словораздела и зона его правого контекста. Каждая зона имеет в своем составе определенный набор характеристик (признаков разного типа). Текстовая зона представляет собой пословную орфографию целевого текста, где каждое графическое слово занимает отдельную строку² таблицы (БДС), номер которой соответствует порядковому номеру словораздела, следующего за словом с тем же порядковым номером. Зона левого контекста в настоящее время включает только частеречный признак слова³, стоящего непосредственно перед словоразделом; планируется ввести также информацию о типе объемлющей синтаксической составляющей, граничащей слева с данным словоразделом. Кроме того, фиксируется несколько фонетических признаков: длительность/длина речевого фрагмента перед словоразделом. Здесь же приводятся акустические показатели звуковых сегментов конечного участка просодической составляющей перед словоразделом. Центральная зона словораздела включает информацию о знаке препинания на словоразделе, формальные синтаксические «скобочные» признаки – общее количество скобок НС, соотношение закрывающих и открывающих скобок. Из фонетических

¹ С точки зрения реальных механизмов, контролирующих появление ПШ в тексте, естественно предполагать, что диктор при озвучивании текста ориентируется в известной степени на знаки препинания, а также и на стоящий за ними синтаксис.

² Для упрощения БД некоторые последовательности из «мелких» служебных слов с союзным, модальным, наречным значением условно считались одним графическим словом, т.е. устойчивым выражением, эквивалентным слову (УВР), (*будто бы, потому что, в течение, или же и др.*

³ Базовый список различаемых частей речи был взят из Национального Корпуса Русского Языка 2003-2005.

признаков – длительность паузы и наличие в ней включенного вдоха. Состав признаков зоны правого контекста словораздела, зеркально аналогичен составу зоны левого контекста.

Таблица, приведенная в конце доклада, представляет собой сокращенный вариант полного макета: в ней показаны только те атрибуты текущей версии БДС, которые по результатам статистической обработки данных оказались наиболее значимыми для моделирования ПЧ. Например, из 12-ти факторов синтактико-пунктуационной разметки показаны только пять, а из разметки по методу Гаспарова-Скулачевой только основной тип связи, имеющий максимальную предикторную силу для выбора типа и локализации ПШ.

Что касается электронного формата БДС, то он ограничен рабочей книгой EXCEL и рабочей таблицей статистического пакета, который допускает более красивое оформление и более удобен для построения математических моделей ПЧ с использованием *вектора* текстовых признаков словораздела, имеющих различную лингвистическую природу. Оба формата могут использоваться также для получения «линейных» орфотекстов с разными типами разметки (паузальной, ПШ, синтаксической, частеречной и пр.).

В заключение следует сказать, что созданная нами база должна быть расширена в нескольких направлениях. Так, оставаясь в рамках того же материала, в фонетическую зону БДС нужно включить информацию о наличии, типе и степени выделенности тонального акцента на словах, разделенных словоразделом, а в зону собственно словораздела – информацию о наличии в словоразделе особых фонетических явлений: вдохов разного типа и глубины, признаков ларингализации и, возможно, других особенностей фонации. При наличии нескольких прочтений одного и того же текста одним и тем же или разными дикторами, с разными установками на степень выразительности чтения, нужно предусмотреть возможность оперативного и удобного включения в БДС характеристик этих прочтений и их статистическую обработку. Это позволит уточнить понятие и содержание разграничения между обязательным и факультативным ПЧ, которое достаточно часто декларируется в интонационной литературе. Наконец, оставаясь в рамках предложенной структуры БДС, нужно предусмотреть и опробовать возможности введения и анализа информации о ПЧ в других типах монологического дискурса: поэтической, публичной, научной, театральной, спонтанной речи с разной степенью подготовленности и т.д. В методическом плане необходимы дальнейшие исследования зависимости ПЧ от вектора признаков словораздела методами многомерной классификации/регрессии с целью построения формальных моделей адекватного прогноза ПЧ: локализации, глубины ПШ и их фонетической реализации. Такие модели необходимы не только для расширения фонетических знаний в области фразовой просодии, но и для создания компьютерных систем качественного синтеза и распознавания речи в разных дискурсивных контекстах¹.

¹ В приводимой ниже табличной иллюстрации БДС факторы, наиболее значимые для выбора локализации и глубины ПШ, выделены жирным шрифтом.

База дискурсивных признаков словораздела в устной русской речи (фрагмент)

Номер словораздела	Текст ИГ в по- словной орфографи- и и одновременно левый контекст словораздела	Знак Прелиминация	ЗП+	ЗП 0/1	ЗП 0/1/2/3	Число закрывающих скобок	Общее число откр. скобок на словоразделе	Всего скобок на словоразделе	Число следующего предложения	ПШ по согласованной группе аудиторгов	А б з а н	Длительность паузы (мс)	Длительность речевого фраг- мента (мс)	Крайние сегменты	Часть речи левый контекст (ЛК)	Часть речи правый контекст (ПК)	Гр-ца финитной клаузы	Бессоюзное сочинение	Союзное сочинение	Поянительный оборот	Левая граница в обороте	Иерархический тип синтаксической связи
0						0	2	2	2							сущ						
1	Люда		-	0	0	0	1	1	0	0	0	0,0		-	сущ	сущ	0	0	0	0	0	а
2	Величко		-	0	0	2	0	2	0	1	0	128,0	816,0	-	сущ	глагол	0	0	0	0	0	р
3	родилась		-	0	0	0	1	1	0	0	0	0,0		-	глагол	союз	0	0	1	0	0	с
4	и		-	0	0	0	1	1	0	0	0	0,0		-	союз	глагол	0	0	0	0	0	а
5	росла		-	0	0	0	1	1	0	0	0	0,0		-	глагол	предл	0	0	0	0	0	т
6	в		-	0	0	0	1	1	0	0	0	0,0		-	предл	сущ	0	0	0	0	0	а
7	провинции	,	зпт	1	1	1	1	2	0	2	0	540,0	1615,0	-	сущ	предл	0	0	0	1	1	сов
8	в		-	0	0	0	1	1	0	0	0	0,0		-	предл	числ	0	0	0	0	0	а
9	одном		-	0	0	0	1	1	0	0	0	0,0		-	числ	предл	0	0	0	0	0	і
10	из		-	0	0	0	2	2	0	0	0	0,0		-	предл	прил	0	0	0	0	0	а
11	среднерусских		-	0	0	1	1	2	0	1	0	0,0		-	прил	прил	0	0	0	0	0	а
12	неперспективных		-	0	0	1	0	1	0	0	0	0,0		-	прил	сущ	0	0	0	0	0	а
13	городков	.	тчк	1	3	8	2	10	2	4	0	1114,0	2900,0	-	сущ	числ	1	0	0	0	0	f
14	Одна		-	0	0	1	0	1	0	0	0	0,0		-	числ	сущ	0	0	0	0	0	а
15	фабрика,	,	зпт	1	1	0	2	2	0	2	0	353,0	858,0	-	сущ	прил	1	1	0	0	0	с
16	лесопильный		-	0	0	1	0	1	0	0	0	0,0		-	прил	сущ	0	0	0	0	0	а
17	завод	,	зпт	1	1	0	2	2	0	2	0	358,0	1109,0	-	сущ	прил	1	1	0	0	0	с

115	пиявки.	.	тчк	1	3	2	2	4	2	5	1	3174,0	2316,0	-	сущ	прил	1	0	0	0	0	f

ЛИТЕРАТУРА

1. *Ladd, R.* Intonational phrasing: a case of recursive prosodic structure / R. Ladd // *Phonology*. – Yearbook (Vol) 3. – P. 311–340.
2. *Ladd B.* Theories of prosodic structure: evidence from syllable duration / B. Ladd, D. R. Campbell // *Proc. of the 12th Congress of Phonetic Sciences, Aix-en-Provence, France*. – 1991. – P. 290–293.
3. *Sanderman, A.* Prosodic Phrasing (production, perception, acceptability and comprehension) / A. Sanderman. – Eindhoven, 1996. – 138 p.
4. *Selkirk, E.* Phonology and syntax: the relation between sound and structure / E. Selkirk. – Cambridge : MIT, 1984. – 476 p.
5. *Аванесов, Р. И.* Русское литературное произношение / Р. И. Аванесов. – М. : Просвещение, 1972. – 383 с.
6. *Кривнова, О. Ф.* Интонационное членение и сегментирующая сила словоразделов в звучащем тексте (данные перцептивного эксперимента) / О. Ф. Кривнова, С. В. Князев, О. С. Смирнова // *Труды Института русского языка им. В. В. Виноградова*. – Выпуск 17. Фонетика. – М., 2018. – С. 128–139.
7. *Смирнова, О. С.* Статистический анализ результатов перцептивного оценивания глубины просодических швов в русском звучащем тексте // *Компьютерная лингвистика и интеллектуальные технологии : По материалам ежегодной международной конференции «Диалог» (Москва, 31 мая – 3 июня 2017 г.)* : в 2 т. – М. : Изд-во РГГУ, 2017. – С. 12. – [Электронный ресурс]. – Режим доступа : URL: <http://www.dialog-21.ru/media/3980/smirnovaos.pdf>
8. *Гаспаров, М. Л.* Статьи о лингвистике стиха / М. Л. Гаспаров, Т. В. Скулачева. – М. : Языки славянской культуры, ИРЯ РАН, 2005. – 288 с.