

# ПРОСОДИЧЕСКИЙ ПОРТРЕТ ГОВОРЯЩЕГО КАК ИНСТРУМЕНТ ТРАНСКРИБИРОВАНИЯ УСТНОГО ДИСКУРСА<sup>1</sup> SPEAKER'S PROSODIC PORTRAIT AS A TOOL OF SPOKEN DISCOURSE TRANSCRIPTION

Кибрик А.А. (kibrik@comtv.ru)  
Институт языкознания РАН

В докладе предлагается методологический аппарат, повышающий качество транскрибирования устного дискурса при создании корпусов звучащей речи. Просодические звукоотпечатки, лежащие в основе сегментации дискурса и выражения фазовых значений, идентифицируются при помощи просодических портретов индивидуальных говорящих.

## 1. Вводные замечания

В настоящее время появляется все больше корпусов устной речи на разных языках. Это важный и позитивный процесс – он связан с признанием того факта, что устная форма языка как минимум не менее важна, чем письменная. При этом создание устных корпусов сопряжено с большими трудностями, ведь главный компонент устного корпуса – это даже не собственно звук, а *транскрипт* звука. Транскрибирование устного дискурса, то есть преобразование звучащего дискурса в транскрипт, включает в себя множество решений, которые должны быть систематическими, последовательными и универсальными. Если транскрибирование осуществляется *ad hoc*, без надлежащей теоретической базы, то ценность получаемого продукта невелика, этот продукт не отражает существенных свойств исходного объекта, то есть устного дискурса.

Данный доклад основан на опыте разработки транскрипции устного дискурса, связанном с подготовкой устного русского корпуса «Рассказы о сновидениях», см. Кибрик и Подлеская 2003. Одна из задач этого проекта состоит в том, чтобы отразить специфику организации устного дискурса при помощи дискурсивной транскрипции.

## 2. Сегментация и фаза

К числу важнейших компонентов дискурсивной транскрипции относятся сегментация дискурса и выражение фазовых значений (см. Кибрик и Подлеская 2006). Дискурс продуцируется говорящим не в виде плавного потока, а в виде квантов или сегментов – элементарных дискурсивных единиц (ЭДЕ). ЭДЕ идентифицируются при помощи комплекса просодических параметров, включая паузацию, темп, громкость, наличие единого тонального контура и главного акцента. В литературе по просодической сегментации устной речи чаще используются другие термины – фраза, интонационная фраза, интонационная группа, ритмическая группа, интонационная единица (см., например, Светозарова и др. 1988; Chafe 1994: 57; Хитина 2004; Кривнова 2007: раздел 2.3).

Как правило, каждая ЭДЕ маркирована с точки зрения *фазы* (термин из работы Кодзасов 2002), то есть по признаку конечности/неконечности. Можно различать разные иерархические уровни фазовых значений. Так, иллокутивная фаза связана с конечностью/неконечностью в коммуникативном взаимодействии: общий вопрос является неконечной иллокуцией, сообщение – конечной. В рамках одной иллокуции ЭДЕ также квалифицируются как конечные или неконечные; этот тип фазы можно назвать внутренним. Основным средством маркирования фазы в устном дискурсе является движение тона в главном акценте (см. Кодзасов 2002). Говоря наиболее обобщенно, конечная иллокутивная фаза маркируется нисходящим (падающим) тоном, неконечная иллокутивная – восходящим, а неконечные внутренние ЭДЕ зеркально адаптируются к конечной.

В дискурсивной транскрипции сегментация на ЭДЕ обычно передается делением на строки, а фаза – при помощи пунктуационных знаков в конце строк. Пунктуационные знаки в этом случае используются не просто как аналоги пунктуации, принятой в письменной форме языка, а как строгие обозначения, расставляемые на основе дискурсивной семантики и просодических фактов. Именно такая система обозначений принята в транскрипции, разрабатываемой в проекте «Рассказы о сновидениях», см. Кибрик и Подлеская 2003, 2009.

<sup>1</sup> Данное исследование выполнено при поддержке гранта РГНФ 08-04-00165а.

И для сегментации, и для разметки фазы большое значение имеют частотные характеристики, то есть характеристики основного тона голоса говорящего. Приведу два примера. Во-первых, при решении вопроса о границе ЭДЕ используется понятие нейтрального тонального уровня, с которого говорящий начинает обычную ЭДЕ; выход на такой нейтральный уровень иногда называют «ресет». Наиболее типичный тональный паттерн состоит в том, что после начала с нейтрального уровня происходит подъем тона, а к концу деклинация. Во-вторых, при решении вопроса о конечности/неконечности данной ЭДЕ в рамках иллокуции используется представление о самом низком уровне тона, характерном для данного говорящего. Дело в том, что в речи большинства носителей русского языка имеются два семиотически противопоставленные типа падающего тонального акцента:

- конечный – так называемая интонация точки: падение в нижний уровень голоса, комфортный для данного говорящего
- неконечный – так называемая интонация запятой с падением: падение в уровень, на 2-4 полутона выше нижнего

### 3. Относительность просодических характеристик

Таким образом, оба рассмотренные просодические явления – ресет и тип падения – являются не абсолютными, а относительными, зависят от характеристик голоса индивидуального говорящего. Если транскрайбер (то есть эксперт, выполняющий транскрибирование устного дискурса) не располагает информацией о голосе данного говорящего, он не может принять решение о том, является ли уровень тона в начале данной ЭДЕ ресетом, и о том, является ли данное падение конечным или неконечным.

Это напоминает общеизвестную ситуацию, связанную с различием между фонетикой и фонологией. Когда носитель языка сталкивается с конкретным звуковым сегментом, в огромном большинстве случаев он в состоянии идентифицировать эту конкретную инстанцию как экземпляр фонемы. Как именно эта идентификация происходит – предмет научного исследования. Одни носители шепелявят, другие грассируют, третьи склонны к фарингализации и т.д. Реализация фонем чрезвычайно вариативна по говорящим и – в рамках одного говорящего – по различным фонетическим контекстам, по функциональным стилям, степени формальности речи и многим другим параметрами. Тем не менее, нормальный носитель языка справляется с этой задачей без больших усилий. Лингвисту, даже имеющему в своем распоряжении арсенал акустической аппаратуры, не так легко смоделировать этот процесс распознавания. Однако ему помогает его собственная интуиция как носителя и достаточно надежные знания о составе фонем языка и о возмущающих факторах.

Те же проблемы имеются в области несегментных аспектов звука – просодии. Носитель языка, конечно, легко распознает все значимые просодические феномены и понимает их семантику. А положение лингвиста здесь заметно сложнее. Во-первых, сами просодические звукотипы известны менее надежно, чем в сегментной сфере. Во-вторых, отсутствует очевидный метаязык, который описывал бы особенности конкретных говорящих. Данный доклад связан с последним обстоятельством. Прежде чем принимать решения о сегментации и расстановке пунктуационных знаков в транскрипте дискурса данного говорящего, необходимо вначале изучить просодическую систему этого человека. Описание наиболее важных элементов этой системы можно назвать *просодическим портретом*. Можно сказать, что рассматриваемая здесь задача – необходимость увидеть просодическую фонологию за просодической фонетикой.

### 4. Компоненты просодического портрета

С точки зрения принятия решений о сегментации и об отражении фазовых значений при работе над корпусом «Рассказы о сновидениях» наиболее важными оказались следующие элементы просодического портрета говорящего:

- тональный диапазон
- типичный тон начал (ресетов)
- целевой уровень конечных падений
- целевой уровень неконечных падений
- целевой уровень подъемов при запятыях
- уровень заударного падения при запятыях
- уровень подъема при многоточиях

Тональный диапазон характеризует голос говорящего в наиболее грубой форме. Этот диапазон показывает минимальные и максимальные значения частот, которые голос говорящего принимает на протяжении дискурса, имеющегося в распоряжении исследователя.

Типичный тон начал (ресетов) – это нейтральный уровень, с которого говорящему удобно начинать каждую очередную ЭДЕ. Этот тон представляет собой не единичное значение, а определенную полосу. Иначе говоря, он варьирует. Однако это варьирование жестко ограничено. Разница между максимальным и минимальным значениями находится в пределах нескольких полутонов.

Целевой уровень финальных падений – это еще один базовый для говорящего уровень частоты, в который стремится упасть его голос в конце иллокутивного акта сообщения. По мнению О.Ф. Кривновой (личное сообщение), целевой уровень финальных падений совпадает с абсолютным нижним уровнем голоса данного говорящего. Целевой уровень финальных падений также представляет собой некоторую полосу.

В отличие от финальных падений, нефинальные падения происходят в уровень частоты, чуть более высокий. Полоса нефинальных падений иногда четко отличается от полосы финальных падений. Иногда, правда, эти полосы могут пересекаться, подобно тому как реализации некоторых фонем могут подвергаться нейтрализации.

При канонической интонации запятой подъем также происходит в некоторую фиксированную полосу. Размер этой полосы может быть большим: некоторые говорящие временами срываются на фальцет, и тогда интонация запятой может реализовываться очень высоким подъемом тона.

В русской речи (и этим она отличается от многих других европейских языков) при интонации запятой за семиотически значимым подъемом следует автоматическое падение тона. Это падение представлено тогда, когда несущий акцент не приходится на последний слог ЭДЕ; в последнем случае автоматическое падение, как правило, отсутствует. В тех случаях, когда падение есть, представляет интерес его целевая полоса.

Наряду с интонацией точки и запятой, очень часто в корпусе встречается «интонация многоточия». В этом случае чаще всего происходит подъем в уровень, значимо более низкий, чем при канонической интонации запятой. Передаваемая семантика может быть наиболее обобщенно описана как неопределенность фазовой характеристики: говорящий не может определить, является ли ЭДЕ конечной или нет. Значимость этого просодического феномена для исследуемого нами жанра дискурса столь велика, что его также необходимо включить в просодический портрет. При этом опять же оценивается целевая полоса подъема тона при наиболее типичных реализациях многоточия.

Кроме того, в просодический портрет включается пункт «прочие просодические характеристики», где фиксируются различные особенности, которые потенциально могут оказаться важными для того или иного аспекта транскрипции.

### **5. Примеры просодических портретов**

Рассмотрим несколько примеров просодических портретов конкретных рассказчиков. Отметим, что все числовые значения приводятся в герцах (Гц).

тональный диапазон	типичный тон начал (ресетов)	целевой уровень финальных падений	целевой уровень нефинальных падений	целевой уровень подъемов при запятых	уровень заударного падения при запятых	уровень подъема при многоточиях	другие характеристики
160-360	200-220	160-180	190-260	290-360	230-300	210	фрикативное [г]

*Таблица 1. Просодический портрет рассказчицы рассказа Z52 (МК ж 16 лет)<sup>2</sup>*

Как можно видеть, у данной рассказчицы полоса начал (ресетов) очень узка – она составляет всего лишь 20 Гц. Целевые уровни финальных и нефинальных падений различаются достаточно четко. То же касается различия между подъемом при интонации запятой и подъемом при интонации многоточия. Уровень заударного падения при запятых близок к целевому уровню нефинальных падений.

Эти выводы относительно данной рассказчицы имеют лишь ограниченную надежность, поскольку в корпусе имеется только один принадлежащий ей рассказ. Рассмотрим теперь просодические портреты, основанные на шести рассказах одного и того же рассказчика.

<sup>2</sup> Номера рассказов, в соответствии с нумерацией в текущей версии корпуса, состоят из литеры (например, Z) и условного порядкового номера. Идентичность рассказчиков обозначена при помощи специальных кодов. Код состоит из инициалов, указания на пол и возраст.

тональный диапазон	типичный тон начал (ресетов)	целевой уровень финальных падений	целевой уровень нефинальных падений	целевой уровень подъемов при запя- тых	уровень заударного падения при запя- тых	уровень подъема при многогочиях	другие характери- стики
190 – 500+	240-290 (первый 310)	190-200	220, часто с загибом	400-500, однажды 300	220-240	280-360	Хорошо выраженные дви- жения тона.. Упередненная артикуляция. Говорит немного в нос.

Таблица 2. Просодический портрет рассказчика рассказа Z35 (АМ м 9 лет)

тональный диапазон	типичный тон начал (ресетов)	целевой уровень финальных падений	целевой уровень нефинальных падений	целевой уровень подъемов при запя- тых	уровень заударного падения при запя- тых	уровень подъема при многогочиях	другие характери- стики
210-480	240-280	200	220-240	370-470	220-240	260-340	Хорошо выраженные движения тона.. Упередненная артикуляция. Говорит немного в нос. Часто скрипучий голос на гласных. Почти все идет на многогочиях

Таблица 3. Просодический портрет рассказчика рассказа Z36 (АМ м 9 лет)

тональный диапазон	типичный тон начал (ресетов)	целевой уровень финальных падений	целевой уровень нефинальных падений	целевой уровень подъемов при запя- тых	уровень заударного падения при запя- тых	уровень подъема при многогочиях	другие характери- стики
200 – 500+	280-320	200?	210-250	360-480, однажды 330	230-240	—	В этом рассказе странно высокий уровень ресетов по сравнению с другими у этого рассказчика

Таблица 4. Просодический портрет рассказчика рассказа Z38 (АМ м 9 лет)

тональный диапазон	типичный тон начал (ресетов)	целевой уровень финальных падений	целевой уровень нефинальных падений	целевой уровень подъемов при запя- тых	уровень заударного падения при запя- тых	уровень подъема при многогочиях	другие характери- стики
220-400	260-290	надежных дан- ных нет	240-250	330-400	260-270	290-330 (строки 6-7 – 400Гц?)	Хриплый голос

Таблица 5. Просодический портрет рассказчика рассказа Z39 (АМ м 9 лет)

тональный диапазон	типичный тон начал (ресетов)	целевой уровень финальных падений	целевой уровень нефинальных падений	целевой уровень подъемов при запятых	уровень заударного падения при запятых	уровень подъема при многоточиях	другие характеристики
220-400	250-290	данных нет	220	330-400	240-300	–	Часто в паузах шмыгает носом. К концу уровень подъемов в запятых заметно снижается: вначале около 400 Гц, в конце 340, и к тому же тихо.

Таблица 6. Просодический портрет рассказчика рассказа Z40 (АМ м 9 лет)

тональный диапазон	типичный тон начал (ресетов)	целевой уровень финальных падений	целевой уровень нефинальных падений	целевой уровень подъемов при запятых	уровень заударного падения при запятых	уровень подъема при многоточиях	другие характеристики
190 – 500+	260-290	данных нет	220	370-410	220-260	340-380, однажды 420	

Таблица 7. Просодический портрет рассказчика рассказа Z41 (АМ м 9 лет)

Учитывая все эти данные, можно построить объединенный просодический портрет рассказчика, гораздо более надежный, чем в случае лишь единичного рассказа.

тональный диапазон	типичный тон начал (ресетов)	целевой уровень финальных падений	целевой уровень нефинальных падений	целевой уровень подъемов при запятых	уровень заударного падения при запятых	уровень подъема при многоточиях	другие характеристики
190 – 500+	240-290, редко до 320	190-200	210-250, иногда с загибом	330-500, редко 300	220-270	260-360, редко до 420	Упередненная артикуляция. Говорит немного в нос, часто хрипло. Часто скрипучий голос на гласных. Часто использует интонацию многоточия

Таблица 8. Суммарный просодический портрет рассказчика АМ (м 9 лет)

Рассказчик АМ имеет большой тональный диапазон. Верхняя часть этого диапазона реализуется в случаях ухода в фальцет. Базовый уровень начал ЭДЕ составляет 240-290 Гц, очень редко выходя за пределы этой полосы. Полоса финального падения узкая и четко фиксированная: 190-200 Гц. Нефинальные падения хорошо отличимы от финальных: 210-250 Гц. В тех случаях, когда различие небольшое, говорящий часто пользуется загибом частотной кривой вверх на заударных слогах. При интонации запятой рассказчик уходит в верхнюю часть голосового диапазона, иногда в фальцет. Целевой уровень заударного падения при запятых очень близок к уровню нефинальных падений. Целевой уровень при многоточиях пересекается с диапазоном подъема при интонации канонической запятой. В таких случаях интонация многоточия отличается от интонации запятой благодаря ряду других признаков – удлинению акцентированной гласной, сложному восходяще-ровному тону, медленному характеру подъема, отсутствию заударного падения в низкий уровень и нек.др.

## 6. Применение просодических портретов

Если бы реализация просодических звукотипов была элементарной, то роль просодических портретов сводилась бы просто к перечню характеристик того или иного говорящего. Это позволило бы решить относительно простую проблему: люди имеют разный тональный диапазон и проводят семиотические различия внутри него по-разному. Например, то, что для одного является типичным целевым уровнем подъема при запятой, для другого является уровнем, характерным для интонации многоточия.

Более сложной является другая проблема. В речи каждого говорящего есть более и менее прототипические инстанции одного и того же просодического звукотипа. Как уже отмечалось выше, семиотически противопоставленные звукотипы в отдельных случаях могут нейтрализоваться. Просодические портреты помогают выявить базовую, исходную семиотическую систему данного говорящего и анализировать более сложные случаи уже на ее основе.

Кроме того, многие просодические паттерны конституируются не одной акустической характеристикой, а целым их набором – ср. список компонентов интонации многоточия в конце раздела 5. В одних случаях оказываются более значимыми одни из этих компонентов, в других – другие. Просодические портреты помогают обнаружить значимые компоненты в каждом отдельном случае и опереться при определении паттерна на твердую эмпирическую основу.

Пример дискурсивной транскрипции (рассказ Z46, рассказчица КЖ ж 14 лет)

11. ... (0.8) /Потом .. (0.2) нам встретился <тут | какой-то> /мостик,
12. ... (0.6) очень .. (0.1) ' Лузенький,
13. ... (0.7) <и> /мы через него еле /проехали тоже,
14. .... (1.2) /потом подъехали к \п~~л~~яжу,
15. ... (0.5) (/Я вообще плавать не \уме-ею.
16. .. (0.4) Не \уме~~л~~а тогда,
17. *когда мне это \сн~~и~~лось.)*
18. .... (1.1) \вот,
19. и-и' .... (1.1) /я почему-то \поплыла.
20. /Нырнула,
21. и \поплыла.

Строки обозначают номера ЭДЕ в рамках рассказа. Случаи интонации точки представлены в строках 15, 17, 19 и 21. Интонация запятой с падением – в строках 12, 14, 16, 18. Обоснованное различие двух типов падений стало возможно в этом рассказе – как и во многих других случаях – лишь при помощи построения просодического портрета данной рассказчицы и выяснения семиотических противопоставлений, используемых ею в частотном континууме. В строках 11, 13 и 20, наконец, можно видеть каноническую интонацию запятой с подъемом тона в главном акценте.

## 7. Заключительные замечания

Несомненно, представленный здесь формат просодического портрета говорящего – лишь самый первый эскиз методологического инструмента. Более детальная разработка этого инструмента – дело дальнейших исследований. К примеру, раздел «прочие характеристики» должен быть расщеплен на ряд значимых позиций, характеризующих разные слои просодии.

Возможный контраргумент против предложенного здесь подхода состоит в том, что транскрибирование устного дискурса оказывается слишком трудоемким процессом. При этом подходе исследователь вынужден осуществлять «нулевой цикл» анализа, который предшествует собственно транскрибированию и на котором изучается система данного говорящего. Возразить на это нечего, кроме того, что такова реальность. Транскрибирование устного дискурса – это действительно трудо- и времяемкий процесс. Затраты усилий отчасти окупаются тем, что получается продукт более высокой ценности.

**Список литературы**

1. Кибрик А.А., Подлеская В.И. 2003. К созданию корпусов устной русской речи: принципы транскрибирования // Научно-техническая информация. Серия 2, 6. С. 5-11.
2. Кибрик А.А., Подлеская В.И. 2006. Проблема сегментации устного дискурса и когнитивная система говорящего // Соловьев В.Д. (ред.) Когнитивные исследования. Вып. 1. М.: Институт психологии РАН. С. 138-158.
3. Кибрик А.А., Подлеская В.И. (ред. 2009.) Семантика и структура устного дискурса (на материале корпуса рассказов о свидениях). М: ЯСЛ (в печати).
4. Кодзасов С.В. 2002. Фазовая символика тона // Арутюнова Н.Д. (ред.) Логический анализ языка. Семантика начала и конца. М.: Индрик.
5. Кривнова О.Ф. 2007. Ритмизация и интонационное членение текста (опыт теоретико-экспериментального исследования). Диссертация на соискание ученой степени доктора филологических наук. М.: МГУ им. М.В. Ломоносова.
6. Светозарова Н.Д., Вольская Н.Б., Павлова А.В., Шитова Л.Ф. 1988. Просодическая организация русской спонтанной речи // Светозарова Н.Д. (ред.) Фонетика спонтанной речи. Л.: Изд-во Ленинградского университета. - 141-182.
7. Хитина М.В. 2004. Делимитативные признаки устно-речевого дискурса. М.: МГЛУ.
8. Chafe Wallace. 1994. Discourse, consciousness, and time. Chicago: University of Chicago Press.