

Инструкция по выполнению вербально-просодической транскрипции (вокальной аннотации)

1. Общая схема работы и содержание инструкции

Основные принципы вербально-просодической разметки, используемой в проекте, были разработаны в коллективной монографии Кибрик, Подлеская (ред.) 2009. В данной инструкции по большей части обсуждаются технические вопросы, а также разбираются аналитические приемы по выполнению разметки. Содержательные вопросы затрагиваются в меньшей степени — в тех случаях, когда это необходимо для объяснения технических решений, а также тогда, когда принципы анализа дискурсивных явлений отличаются от описанных в Кибрик, Подлеская (ред.) 2009. В первую очередь, это касается модификации транскрипционной системы применительно к данным диалогической речи.

1.1. Материалы и программные средства

Для вербально-просодической разметки используются 4 аудиофайла:

- 3 индивидуальных аудиофайла Рассказчика, Комментатора и Пересказчика;
- 1 общий аудиофайл.

Общий аудиофайл используется только как вспомогательный источник анализа. Основная разметка производится на основании индивидуальных аудиофайлов.

Разметка производится отдельно для каждого говорящего. Объединение разметок осуществляется только на уровне отображения результатов разметки (см., в частности, раздел L ниже).

Вся работа выполняется в двух программах:

- текстовый редактор Word (возможна также и работа в аналогичном свободно распространяемом текстовом редакторе, однако оптимальные результаты достигаются именно в Word);
- звуковой анализатор Praat (о работе в Praat см. подробнее раздел 9 данной инструкции).

В Word создаются и редактируются текстовые транскрипты, в Praat — специальные разметочные файлы формата .textgrid. Соответственно, итоговая разметка каждой записи хранится в 6 файлах: трех транскриптах и трех текстгридах.

1.2. Этапы разметки

Аннотирование вербально-просодического компонента проводится в два этапа.

На первом этапе выполняется первичная расшифровка — запись слов, заполненных пауз и некоторых неречевых вокальных действий говорящих (см. раздел 2). Работа ведется в текстовом редакторе.

На втором этапе выполняется точная временная привязка сегментов, деление текстов на ЭДЕ и просодическая разметка (см. разделы 3 – 8). На этом этапе крайне важно обеспечивать синхронизацию между Word и Praat.

Разделение работы на два этапа вызвано следующими соображениями:

- первый этап, не подразумевающей разметки в Praat и не требующий тонкого дискурсивного анализа, можно поручить менее опытным исполнителям; после чего с результатом их разметки могут работать более опытные аннотаторы;
- результаты первого этапа разметки можно использовать для автоматического определения границ слов.

1.3. Автоматическая разметка границ слов

В нашем проекте между первым и вторым этапами разметки производилось автоматическое определение границ слов. Эта работа выполнялась нашими партнерами из компании STEL (<http://speech.stel.ru/>). На вход разработанной STEL процедуры подавались разметки первого этапа, на выходе получался текстовый файл с указанием времени начала и конца произнесения каждого расшифрованного сегмента. В дальнейшем при помощи дополнительного скрипта текстовый файл преобразовывался в формат .textgrid.

Программное определение времени начала и конца произнесения каждого сегмента является удобным, но не обязательным условием разметки. Потенциально эту работу можно проводить вручную, опираясь в целом на принципы, изложенные ниже в разделе 4. Однако при описании этапов разметки в данной инструкции мы исходим из допущения, что автоматические границы слов расставлены.

2. Первичная расшифровка

2.1. Порядок работы

На этом этапе производится первичная расшифровка слов и других вокальных действий трех участников записи. Расшифровка выполняется в текстовом редакторе Word, в отдельном файле для каждого из трех участников. Для прослушивания и анализа файлов рекомендуется использовать программу Praat.

В качестве первого шага мы рекомендуем прослушать целиком общий аудиофайл. Это самый удобный способ ознакомиться с общим содержанием записи, манерой речи говорящих, с местами, содержащими наложения и т.д. **В общем аудиофайле громкость речи каждого говорящего ниже, чем в его индивидуальном аудиофайле, но почти наверняка выше, чем в оставшихся двух файлах.** Для дальнейшей работы по каждому говорящему необходимо использовать уже его индивидуальный файл. Тем не менее, в некоторых записях возможны ситуации, когда слаборазборчивые фрагменты лучше слышны в общем аудиофайле, чем в индивидуальном. Это обстоятельство следует иметь в виду, когда не удастся расслышать фрагмент в индивидуальном аудиофайле.

Примечание. Текст, выделенный красным, относится к записям 2015 года. Для записей 2017 года W-файл представляет собой объединение трех индивидуальных файлов, поэтому в нем громкость каждого говорящего такая же, как в индивидуальных файлах.

В целом идентификация слов и других единиц производится на слух, в сложных местах рекомендуется пользоваться спектрограммой. Результат работы записывается в файл формата .doc. Каждая единица должна быть отделена от другой пробелом или знаком переноса строки. Для удобства восприятия рекомендуется разбивать текст на строки (опытные транскрайберы могут сразу сегментировать текст на элементарные дискурсивные единицы).

2.2. Запись слов

1) Слова записываются в стандартной орфографической форме

Возможны редкие исключения типа *щас*, но в целом они не приветствуются. Если транскрайбер замечает, что слово произносится как-то не стандартно, это следует указать отдельным комментарием.

он пытается их пересчитать, //произносит «пересчиДать»

Также используются конвенциализованная запись для сигналов подтверждения: *угу* и *ага*. В том случае если говорящий произносит эти сигналы не обычным образом, а «полно»

(т.е. как если бы они действительно состояли из звуков, соответствующих этому условному написанию), это имеет смысл указать в примечании:

ага //произносит полно: [ага]

Обязательно использовать букву Ё:

берёт, своём, неё и т.д.

2) При этом все формы слов записываются так, как они произнесены — например, в «неправильных» падежах / числах / временах и т.д. Никакой подгонки материала под то, «как должно быть», не допускается. Например:

и он ээ вторую точнее дополняет,
д= ээ от= **оставшихся** грушами,

3) Также обязательно фиксировать все повторы, в том числе и неполные:

на на свой ба= <какой-то> багажник?
то есть он **пере= пересчитывает** сначала

4) Обязательно записываются оборванные фрагменты слов. Такие фрагменты завершаются знаком равенства без пробела:

он он он какой-то (?) **пож=** он пожилой допустим?

При этом если транскрайбер с достаточной уверенностью полагает, что говорящий планировал произнести какое-то конкретное слово, ему следует записать оборванный фрагмент по правилам стандартной орфографии:

поэтому кажется сначала,
(?) **ч= чт=** что её там нет,

Если же транскрайбер не может с уверенностью понять, что именно оборвано, ему следует записать фрагмент «как слышится»:

Юра правильно сказал (ə) **ш=** про свист мальчиков,

5) Если транскрайбер не уверен, что расслышал именно то слово, которое записал, нужно заключить это слово в угловые скобки:

что Юра **<вот>** как раз рассказал,

Если неуверенность касается последовательности слов, в угловые скобки нужно заключать каждое из слов этой последовательности:

<всё> <там>.

6) Если транскрайбер не может разобрать произнесенное, используется обозначение <НРЗБ>:

потому что он как-то пытается за= сесть на этот **<НРЗБ>**

2.3. Запись звукоподражаний

Звукоподражания заключаются в дизезы и записываются «как слышатся»:

ээ мальчик едет на велосипеде,
#ты-дыш# #ты-дыш# #ты-дыш# #ты-дыш#!

Если звукоподражание сложно записать «как слышится», можно вставить в дизезы описательное выражение:

#рычание#

#блеянье#

Также в дизезы заключаются явления вроде #свиста# (в тех случаях, когда свист используется в качестве аналога лексической единицы).

2.4. Запись заполненных пауз

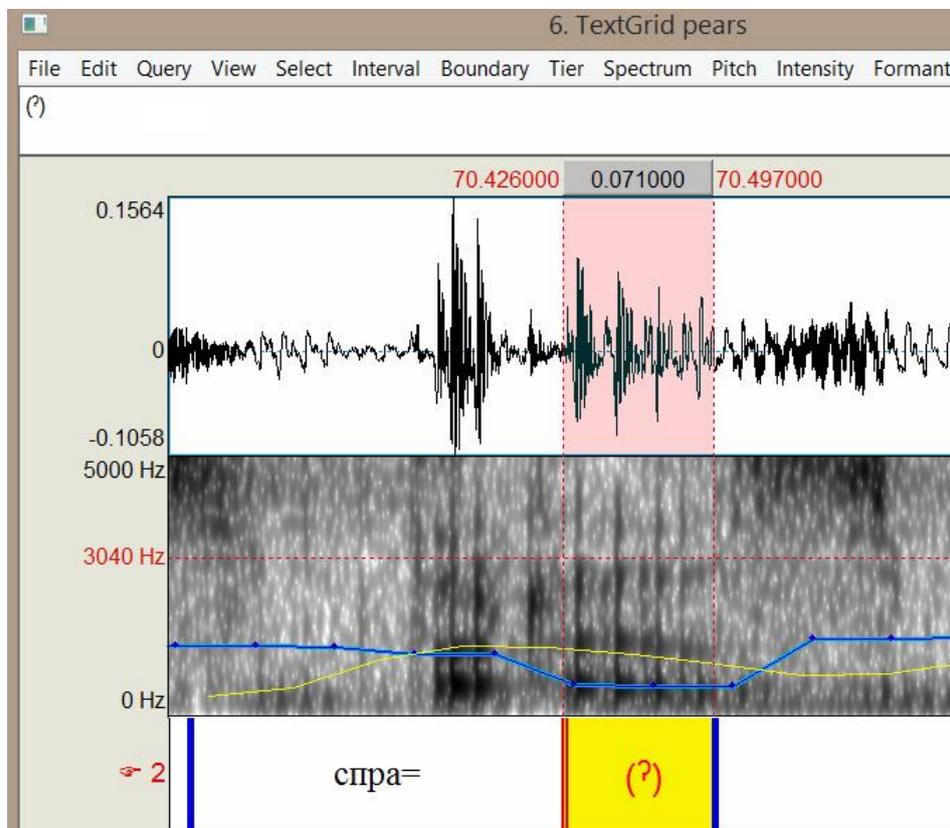
Заполненные паузы — это вокализованные отрезки, которыми говорящий заполняет время хезитации. Каждый говорящий делает это немного по-своему, но мы — с известной долей условности — выделяем следующие 6 типов заполненных пауз:

пауза, заполненная мэканьем	(ш)
пауза, заполненная эканьем	(э)
пауза, заполненная аканьем	(е)
пауза, заполненная гортанной смычкой	(?)
смешанная заполненная пауза, начинающаяся с мэканья	(шэ)
смешанная заполненная пауза, начинающаяся с эканья	(эш)

Как видно из таблицы, все заполненные паузы обозначаются при помощи специальных символов, заключенных в скобки.

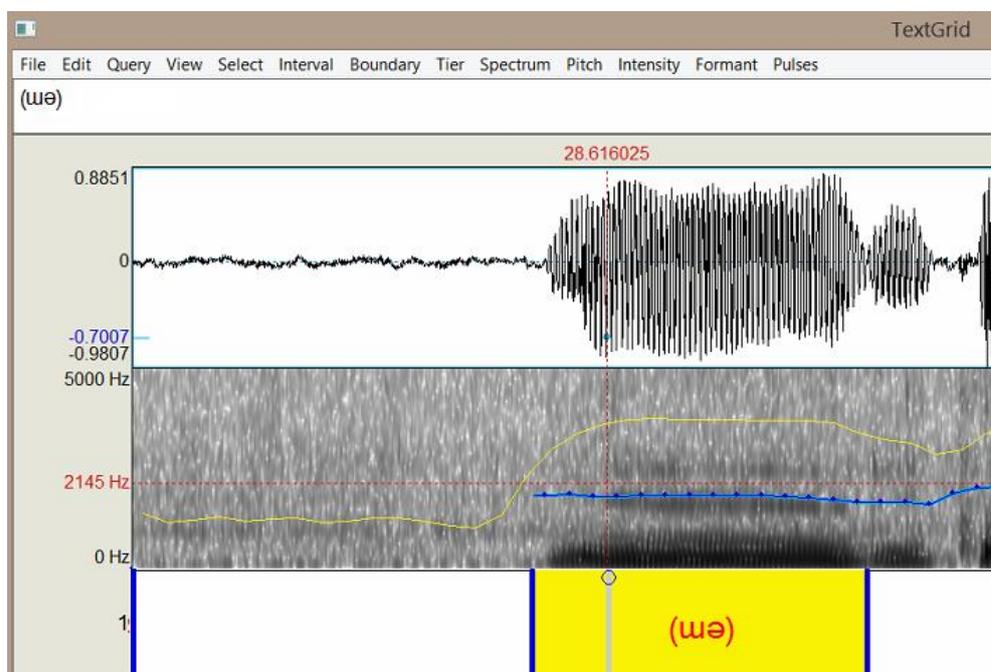
Примечание. На данном этапе работы также допускается использовать другие обозначения: мм, ээ, аа, ", мэ и эм. В дальнейшем они должны быть преобразованы в канонические варианты.

Гомогенные паузы, заполненные простыми звуками, выглядят на спектрограмме обычным для данных звуков образом. Пауза, заполненная гортанным скрипом, выглядит примерно так:



Характерным признаком гортанного скрипа является серия коротких взрывов (вертикальных полос на спектре) и слабая выраженность формант, соответствующих присутствию голоса. На практике существует множество переходных случаев между чистым скрипом (?) и чистым голосом (ə), одним из которых является заполненная пауза (ə), произносимая т.н. скрипучим голосом (см. ниже раздел 5 о явлениях, сопутствующих речи). Для каждого конкретного сомнительного случая необходимо решить (опираясь как на спектрограмму, так и на слух), чего больше: скрипа или голоса. Если больше скрипа, хоть и присутствует голос, следует отмечать паузу как заполненную скрипом (?). Если больше голоса, хотя присутствует и скрип, то следует отметить паузу как (ə), а наличие скрипа отметить в специальном слое Supra аннотационного файла в Praat (см. ниже раздел 5).

Смешанные паузы — это такой способ заполнения паузы, при котором один звук переходит в другой и между разными частями нет никакого промежутка/прерывания. Вот как смешанная пауза выглядит на спектрограмме:



До положения курсора (вертикальный красный пунктир) произносится звук (ш), который дальше переходит в (ə). Пауза типа (əш) выглядит наоборот: длительное эканье оканчивается губной смычкой, длительность которой может быть разной.

2.5. Запись неречевых вокальных действий говорящего

Кроме слов, абсолютных и заполненных пауз, в речи говорящих могут встречаться разнообразные неречевые явления: цоканья языком, чмоканья и пр. Ниже приводится их список с обозначениями.

откашливание	{sp}
цоканье языком	{cl}
чмоканье губами	{sm}
вздых	{sg}
сглатывание	{sw}
свист	{wh}

фырканье	{st}
кашель	{cg}
шмыганье носом	{sf}
хмыканье	{hm}
смех	{laugh}

Вздох отличается от вдоха тем, что состоит из сочетания шумного вдоха и шумного выдоха.

Откашливание — это то, что в повествовании обычно описывается как «гм», «кхм», «эхем», «прочистил горло» и т.д. Это действие, хотя и не являющееся словом или звукоподражанием, имеет определенную семантику.

Кашель, в отличие от откашливания, — физиологическое явление.

Несколько особняком в ряду неречевых явлений стоит **смех**. Смех может как присутствовать в виде отдельного сегмента, так и накладываться на речь говорящего. Если смех присутствует в записи как самостоятельное явление, то его следует записывать наряду со словами и заполненными паузами. Если же смех у говорящего накладывается на конкретное слово или слова, то он не фиксируется в текстовом формате, а аннотируется в специальном слое Supra в программе Praat (см. ниже раздел 5).

3. Совместная работа в Praat и Word

После завершения первого этапа расшифровки, описанного в разделе 2, дальнейшая разметка происходит одновременно в двух средах: текстовом редакторе Word и звуковом анализаторе Praat. При описании дальнейших действий мы исходим из того, что полученные на первом этапе результаты были обработаны программой по автоматическому распознаванию речи. На выходе были получены файлы аннотации Praat в формате .textgrid (далее — текстгриды), в которых в уровне Words указаны все единицы, расшифрованные на первом этапе, и для этих единиц приведено время начала и конца произнесения. Помимо уровня Words, в текстгридах также имеются еще два уровня: EDU (см. подробнее раздел 6) и Supra (см. раздел 5).

Текстовые разметки формата .doc (далее — транскрипты), в свою очередь, преобразованы в табличную форму со следующими столбцами:

- EDU – номер ЭДЕ
- Transcript – слова и прочие сегменты, входящие в ЭДЕ
- Comments – комментарии в свободной форме

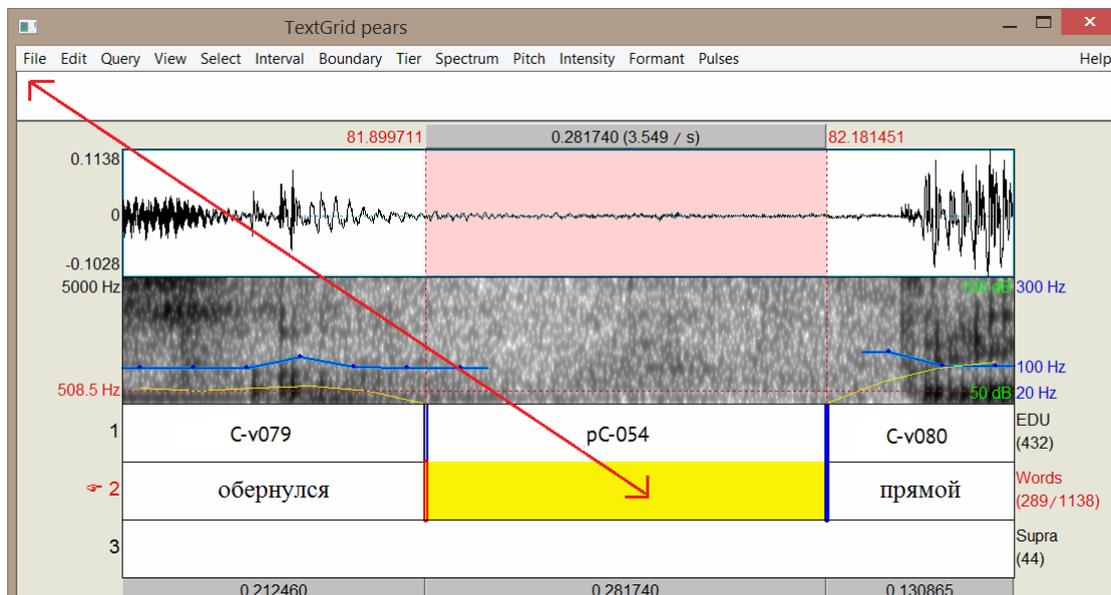
При редактировании разметки крайне важно следить за тем, чтобы изменения в текстгриды и транскрипты вносились согласованно. На каждом этапе разметки необходимо убеждаться, что количество непустых интервалов в слое Words и количество слов и т.д. в столбце Transcript полностью совпадает. В свою очередь, количество пустых интервалов в слое Words должно полностью совпадать с размеченными в транскрипте абсолютными паузами (см. ниже, раздел 4).

Примечание. Для обеспечения согласованности между текстовым транскриптом был написан макрос на языке VBA.

4. Уточнение границ сегментов и запись абсолютных пауз

4.1. Отображение слов и абсолютных пауз

В текстгридах слова, заполненные паузы и прочие явления, описанные выше в разделе 2, представлены в виде непустых интервалов уровня Words, а абсолютные паузы — в виде пустых (неименованных) интервалов этого же уровня:



Важно помнить, что под абсолютной паузой в данном случае понимается отрезок молчания данного говорящего (т.е. другие участники записи в этот момент могут произносить свои реплики). Интервал, соответствующий абсолютной паузе, должен быть абсолютно пустым — он не должен содержать пробелов, разрывов строки и других подобных символов.

В **транскрипте** абсолютные паузы записываются просто пустыми скобками без пробела:

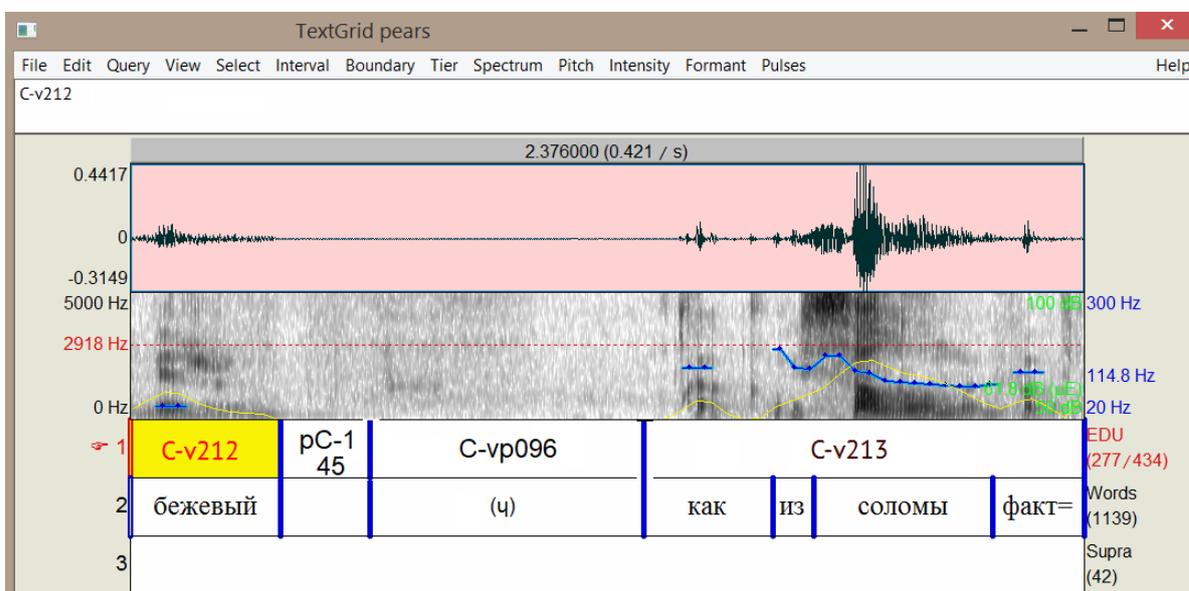
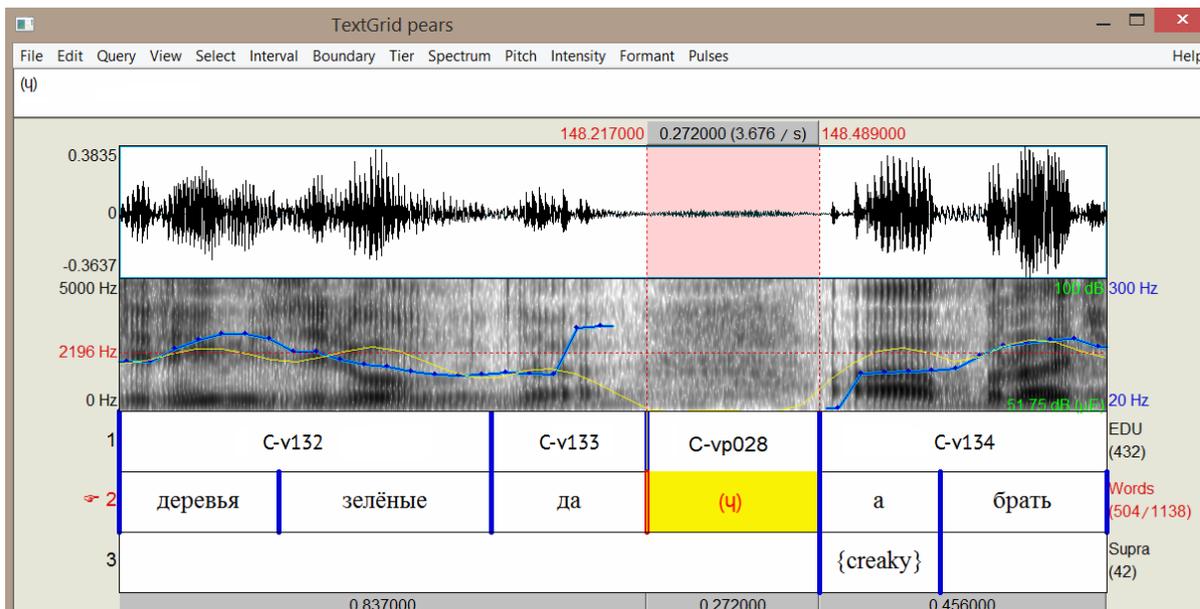
()

Подробнее о месте записи абсолютных пауз в транскрипте см. раздел 6 ниже.

4.2. Вдохи в паузах

В некоторых паузах, не заполненных никакими другими звуками, могут встречаться шумные вдохи, когда говорящий с особым шумом набирает в легкие воздух. **Вдох** следует отличать от **вздоха**: при **вздохе** говорящий шумно набирает воздух, а затем сразу же столь же шумно выдыхает; за шумным вдохом дальше обычно следует речь либо простой выдох.

Паузы, содержащие шумный вдох, обозначаются как (ц). Такие паузы размечаются везде, как в транскрипции, так и в текстгриде. Началом паузы с шумным вдохом считается начало самого шумного вдоха, которое определяется на слух и по спектрограмме. Если до начала вдоха присутствует абсолютная пауза, то она обозначается отдельно, ср. два примера ниже. В первом случае вдох заполняет паузу целиком; во втором — пауза (ц) начинается с момента начала вдоха.



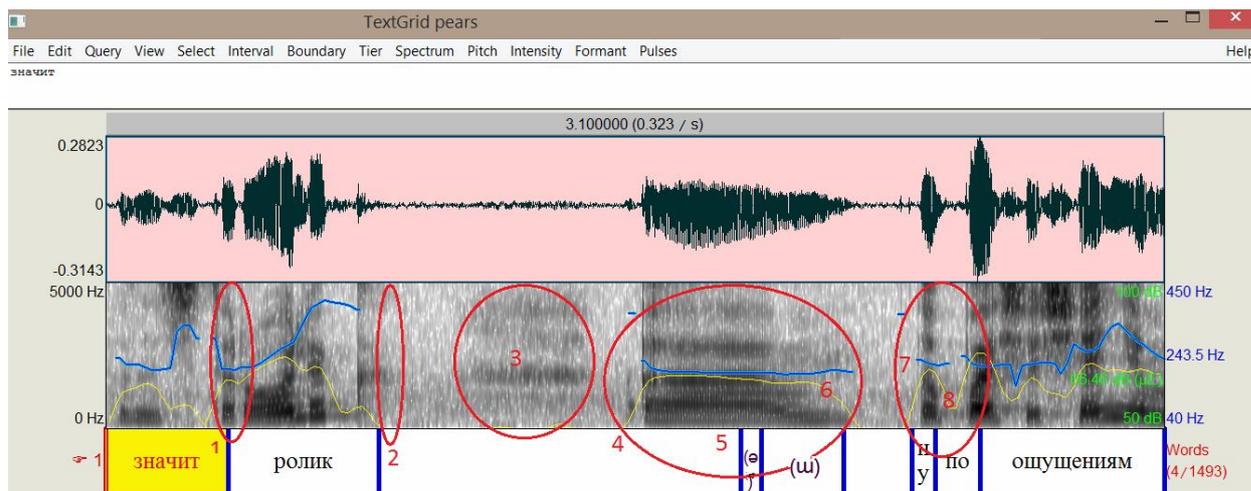
4.3. Проверка правильности интервалов

При определении границ любых интервалов (как пустых, так и содержащих слова, заполненные паузы и другие звуковые явления) можно опираться на:

- слуховое восприятие;
- осциллограмму;
- спектрограмму (Spectrum);
- кривую частоты основного тона (Pitch)
- кривую интенсивности (Intensity).

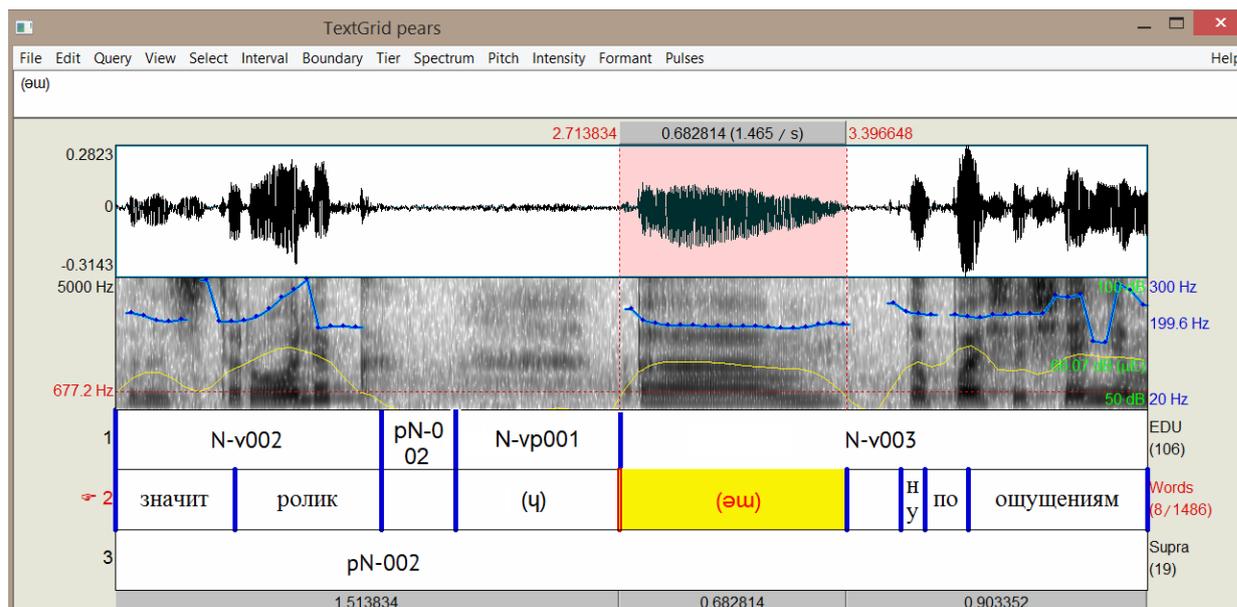
Ни одно из этих средств не может служить абсолютно точным индикатором границы (в частности потому, что запись нескольких говорящих не в студийных условиях всегда содержит множество посторонних звуков, наложения речи двух или более участников и т.д.), однако совокупность нескольких параметров позволяет с достаточной точностью правильно определить место постановки границы.

Возьмем пример из одного файла, ранее обработанного программой автоматической разметки на основе предварительной записи слов.



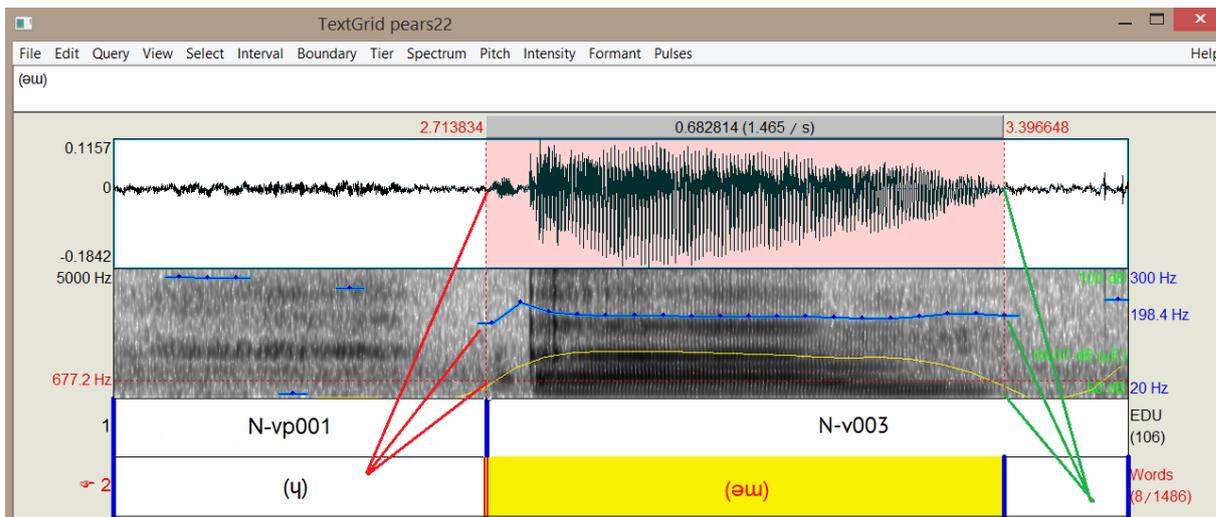
Даже при беглом взгляде заметно, что некоторые явления (например, вдох в центральной паузе) пропущены, а некоторые (следующая за ним заполненная пауза) размечены совсем неправильно. Кроме этого есть еще несколько вызывающих вопросы моментов, которые отмечены на скриншоте цифрами.

Первым делом необходимо исправить явные ошибки (случаи 3, 4 5 со скриншота выше):



Пауза, содержащая шумный вдох, обозначается как (ц) с начала этого вдоха, как он виден на спектрограмме и слышен при прослушивании файла. С этой же границы начинается и ЭДЕ 3 (подробнее об ЭДЕ см. раздел 6 ниже). Далее, вместо двух заполненных пауз «ээ» и «мм» следует проставить одну паузу (эш) и исправить ее границы. В первом варианте значительная часть звука (э) вообще не попала в паузу, а граница между (э) и (ш) не только была лишней, но и проставлена была существенно раньше нужного места.

Посмотрим внимательно на границы заполненной паузы (эш):



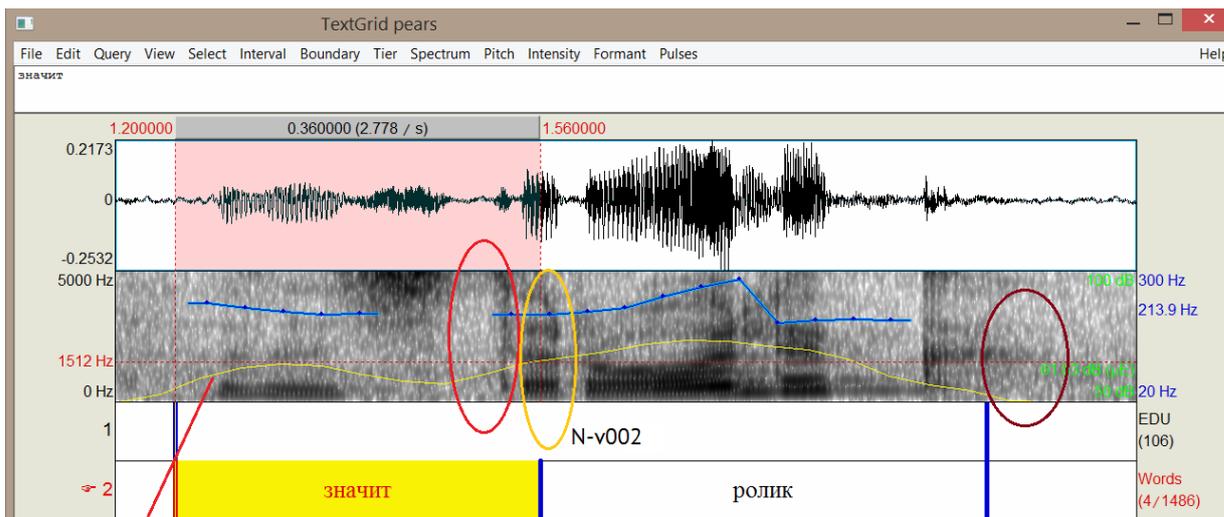
Слева красным обозначены признаки, которые помогают определить начало вокализации:

- увеличение амплитуды звуковой волны на осциллограмме;
- появление первой точки на кривой Pitch;
- появление на спектре первых признаков начала голоса (размыкание голосовых связок);
- резкое увеличение угла кривой интенсивности (что означает более быстрое наращение интенсивности звука — т.е. начала речи).

В совокупности, дополненные слуховой проверкой, эти признаки дают возможность поставить начало (эш) именно здесь, а не, скажем, правее — там, где начинаются форманты (э).

Справа зелеными линиями аналогично отмечены признаки, позволяющие выявить место завершения заполненной паузы.

Вернемся теперь к левому фрагменту (ЭДЕ N-v002) и уточним границы слов «значит» и «ролик».

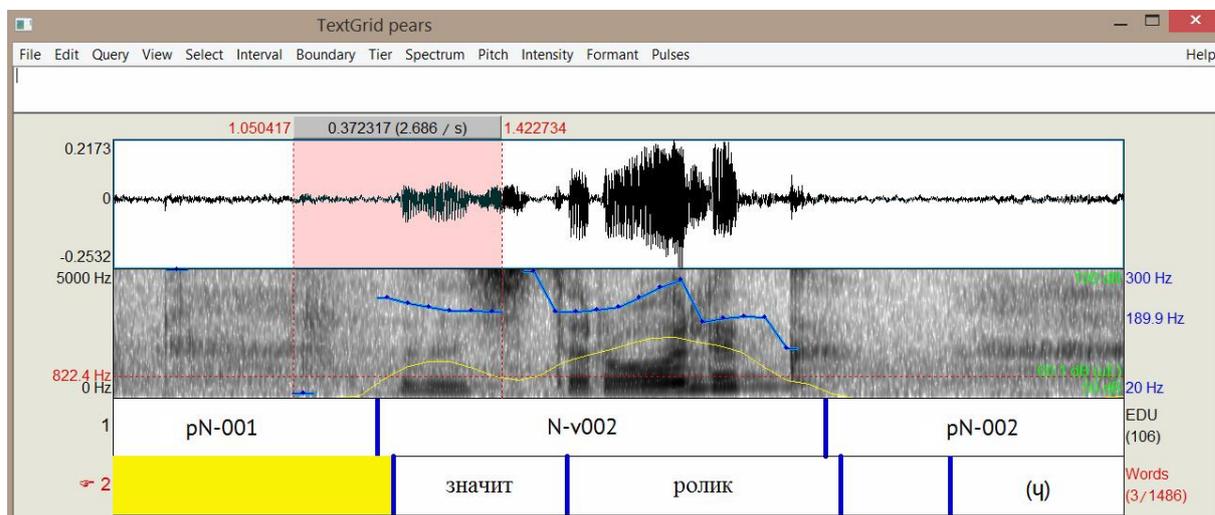


В окне программы можно обнаружить несколько «подозрительных» явлений.

1. Слева начало слова «значит» обозначено существенно левее начала и голоса, и шума на спектре (З - звонкий шумный, значит, должен быть и голос внизу, и шум в верхней части спектра).

2. Граница между словами поставлена не после взрывного Т (красный овал), а на середине Р (желтый овал), что подтверждается еще и при прослушивании.
3. Правая граница слова «ролик» поставлена правее, чем заканчивается придыхание на звуке К (см. бордовый овал).

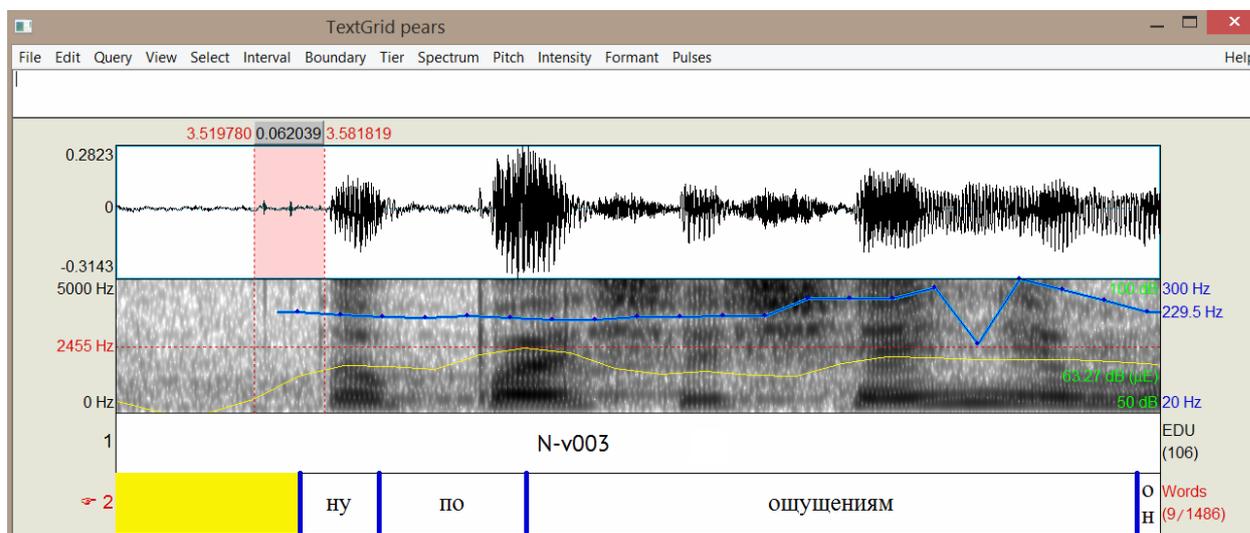
Сдвинем границы слов так, как подсказывает спектр, а затем убедимся, что при прослушивании все правильно (до начала «значит» не должно быть никаких звуков, а окончание К в слове «ролик» должно длиться до новой границы):



Важно понимать, что при проверке границ на слух практически бесполезно прослушивать очень маленькие интервалы, поскольку даже в паузах все равно присутствует небольшой шум, а микрофрагменты звуков меньше определенной длины перестают определяться на слух как фонемы. Поэтому следует прослушивать бóльшие фрагменты до и после проверяемой границы, а затем — фрагмент, в котором потенциальная граница находится примерно посередине, обращая внимание на то, как и в каком месте существенно меняется качество звука.

Например, в рассматриваемом случае для проверки границы «З» лучше прослушивать сначала паузу до «значит» в обоих вариантах, старом и новом, затем само «значит», а затем примерно тот интервал, который выделен на скриншоте выше.

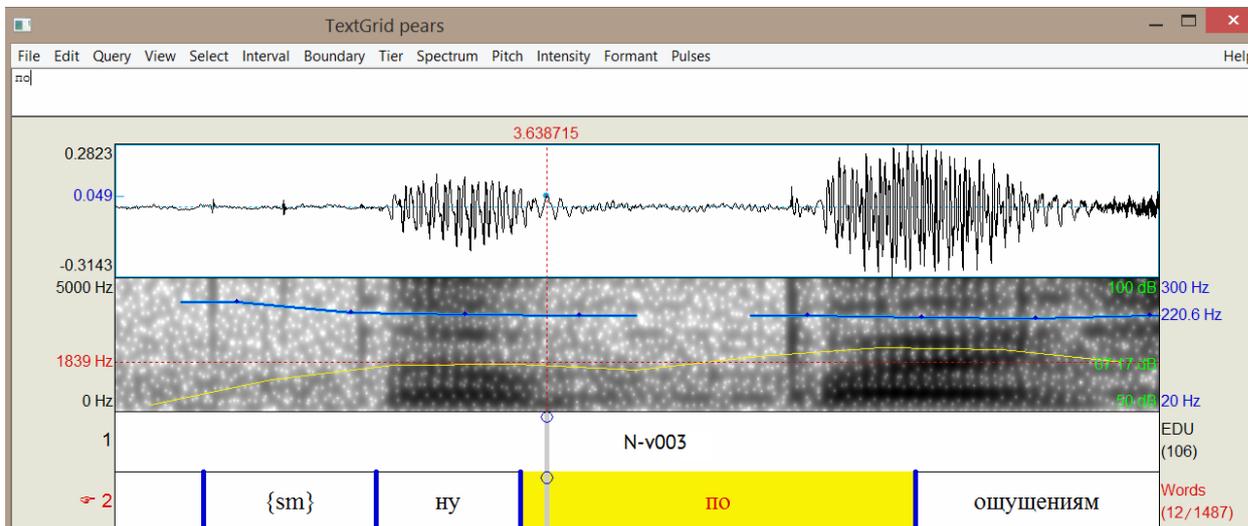
Перейдем к проверке границ в последовательности «ну по ощущениям» (NB! предварительно также необходимо скопировать «новые» границы в слой EDU).



С одной стороны, мы видим, что спектральная картина «ну» (содержащая форманты) начинается позже обозначенной границы; с другой — что в выделенном на скриншоте

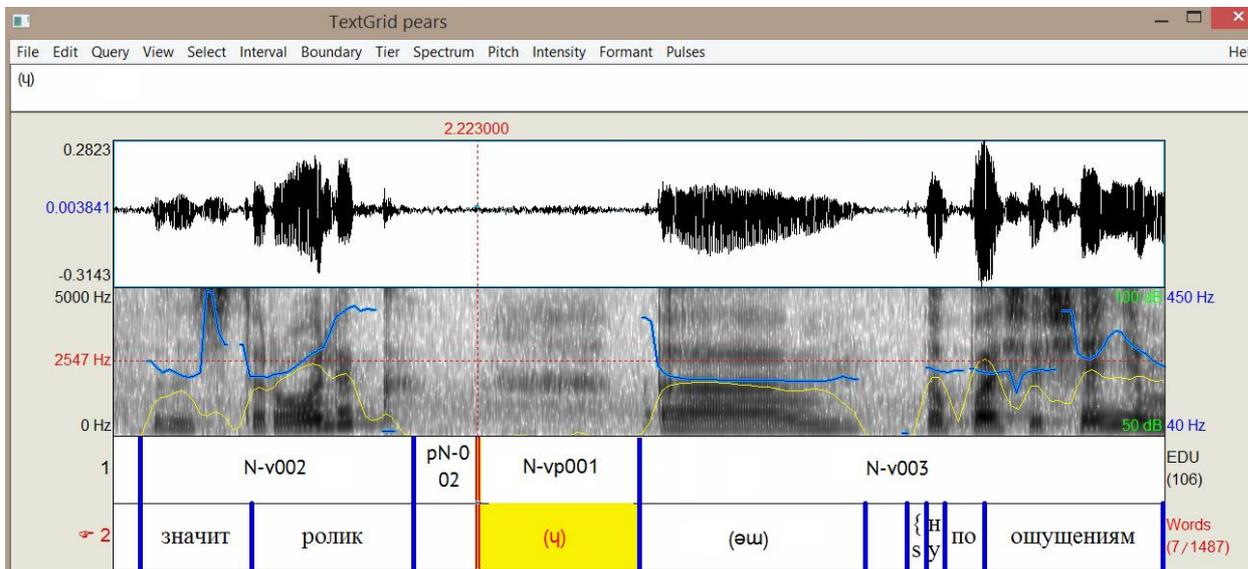
интервале есть три вертикальные полосы, две из которых попали в паузу, а третья попала в границы «ну».

Также видно, что и кривая интенсивности, и кривая Pitch, и осциллограмма подсказывают, что раньше видимого на спектре начала «ну» есть какой-то звук или звуки. Выделив и прослушав этот фрагмент, мы услышим, что это чмокание, с которым говорящая размыкает губы. Чмокание должно быть выделено в свой интервал и отмечено как {sm} в слое Words, а также внесено в транскрипт.



При прослушивании интервала «по», а также при анализе осциллограммы становится видно, что часть затухания звука «у» из «ну» попала в этот интервал. Границу нужно перенести туда, где на скриншоте выше стоит курсор.

В результате всех изменений получаем следующую разметку в текстгриде:



4.4. Некоторые дополнительные соображения

- **Абсолютные паузы между словами** теоретически могут быть любой длительности, даже очень малой, однако чем меньше интервал, тем сложнее доказательно отличить паузу от глухой смычки или другого подобного явления. Поэтому при проверке и разметке границ абсолютной паузы следует придерживаться следующей схемы.

- Паузы протяженностью 0.1 секунды и более выделяются всегда — необходимо только убедиться, что границы звуков до и после паузы отмечены правильно.
- Паузы протяженностью менее 0.1 секунды выделяются только при выполнении следующих условий: (а) пауза определяется как таковая на слух; (б) слева и справа от паузы находятся звуки, границы которых могут быть достоверно определены без сомнений. Это сочетания типа гласный+гласный, гласный+шумный, гласный+сонорный и т.п.
- Если потенциальная пауза менее 0.1 секунды плохо определяется на слух, а слева и справа от нее находятся глухие согласные либо еще какие-то явления затрудняют определение ее границ (например, речь других говорящих, посторонние шумы и т.п.), то пауза **не выделяется**. Межсловная граница в таком случае ставится в самой левой точке, определяемой по спектру и на слух. Таким образом, если между словами и присутствует некий трудно выделяемый микропромежуток, он всегда будет относиться к началу следующего слова.

- Многие служебные слова часто сильно редуцируются. Вместо «потому что» говорящий легко может произносить «пшто» или даже «пш», вместо «то есть» — «тсь» или «сь». При прослушивании всегда обращайте внимание на сильно редуцированные фрагменты. Поскольку программа автоматической разметки ориентируется на стандартную орфографию, почти наверняка границы редуцированных слов и соседних с ними будут размечены неверно.
- В начале и конце слов могут встретиться: губная смычка (w), гласный призвук (ə), гортанная смычка (ʔ), придыхание (ç). Все такие явления, хотя и не отражаются в записи слов в текстгриде, будут позднее обозначены в транскрипте и должны входить в границы слова.

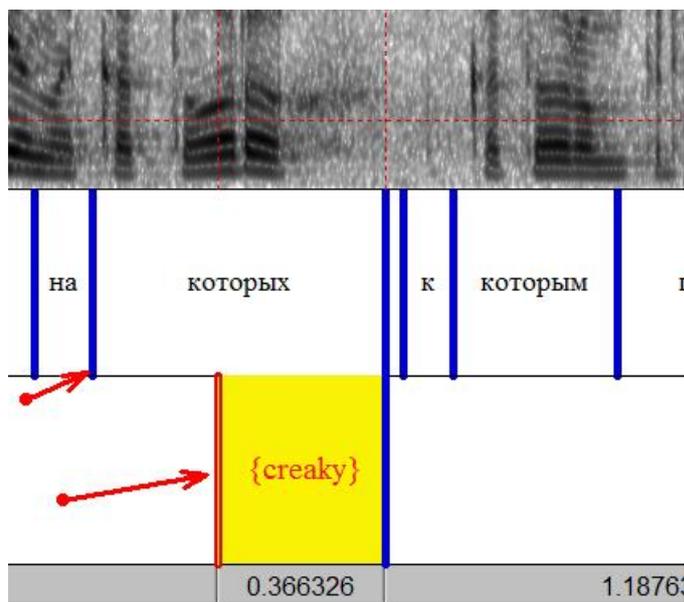
5. Разметка сопутствующих речи явлений

На речь говорящего могут накладываться его же неречевые звуковые действия: смех, улыбка, скрипучий голос. Такие явление нужно размечать в отдельном слое аннотации — Supra.

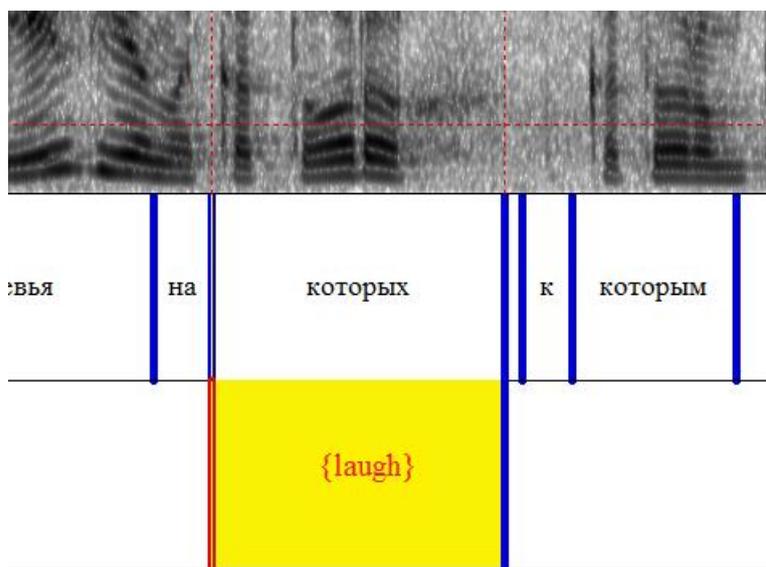
Для нотации неречевых звуковых действий в текстгриде используются следующие обозначения в интервалах:

смех	{ laugh }
улыбка	{ smile }
скрипучий голос	{ creaky }

В целом нужно стараться размечать границы таких сопутствующих явлений независимо от границ интервалов слоя Words: например, если слышно, что скрипучий голос начинается в середине какого-то слова, то это значит, что с начальной частью этого слова интервал скрипучего голоса пересекаться не должен. Выглядит это примерно так:



Но в некоторых случаях установить такие независимые границы невозможно. Например, не всегда легко понять, где начался смех, наложенный на слово. В таких случаях мы условно ставим границу смеха там же, где уже стоит граница слова:



В транскрипте такие явления не отмечаются, но позже могут быть в нее добавлены программным образом.

6. Разметка элементарных дискурсивных единиц

Элементарная дискурсивная единица (ЭДЕ) — это минимальный шаг в построении дискурса. Выделение ЭДЕ — экспертная процедура, основанная на совместной учете коммуникативно-просодических и семантико-синтаксических характеристик. В задачи данной инструкции не входит описание принципов выделения ЭДЕ. Ознакомиться с ними можно, в частности, в Кибрик, Подлесская (ред.) 2009, Коротаяев 2015. Ниже приводятся формальные правила разметки ЭДЕ, а также обсуждаются некоторые нововведения, реализованные в рамках настоящего проекта.

6.1. Границы ЭДЕ в транскриптах

В транскриптах каждой ЭДЕ соответствует одна строка, см. следующий пример, в котором содержится три ЭДЕ:

R-v310	это как {laugh} (в) садовник не \находит одной /корзины,
R-v311	а мальчики едят /груши,
R-v312	проходят \мимо.

Каждую строку с ЭДЕ в транскрипте нужно пронумеровать в столбце EDU. Нумерация должна содержать в себе код говорящего (N, C или R), дефис, литеру “v” (начальная буква слова *vocal*) и трехзначный номер. В окончательном варианте нумерация должна начинаться с 001 и идти без пропусков и повторений. Для предупреждения возможных ошибок и удобства редактирования был написан специальный макрос.

6.2. Границы ЭДЕ в текстгридах

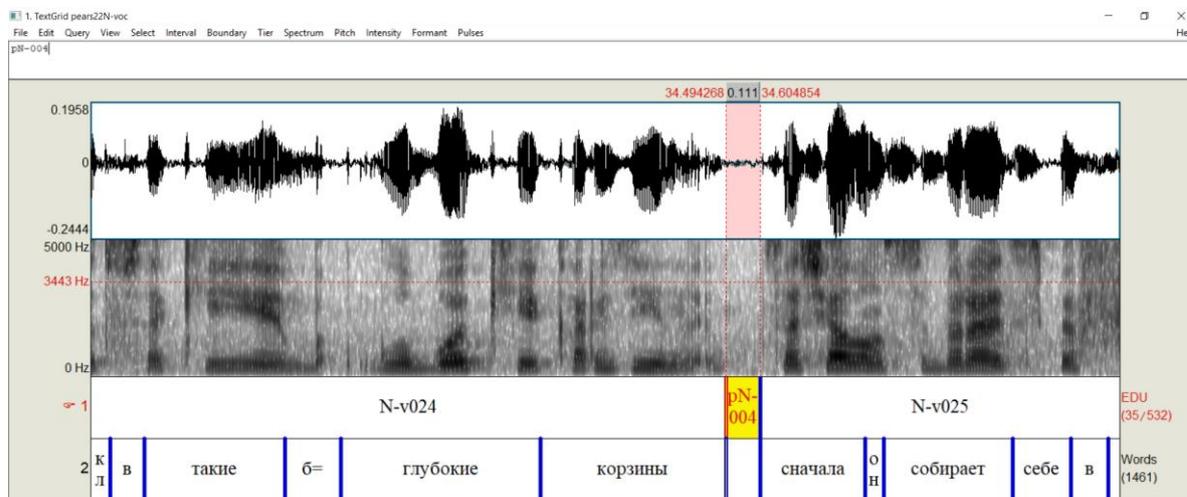
Для аннотации границ ЭДЕ в текстгридах используется уровень EDU.

Все границы интервалов EDU должны полностью совпадать по времени с левыми границами первых слов либо заполненных пауз ЭДЕ и правыми границами последних слов ЭДЕ.

Интервалы, соответствующие ЭДЕ, должны нумероваться точно так же, как в транскриптах.

6.3. Границы ЭДЕ и паузы

1) Абсолютные паузы, располагающиеся между ЭДЕ, **не входят** ни в какие ЭДЕ. В текстгридах в уровне EDU таким паузам соответствуют интервалы с отдельной нумерацией — см. формат на скриншоте и в примере ниже:

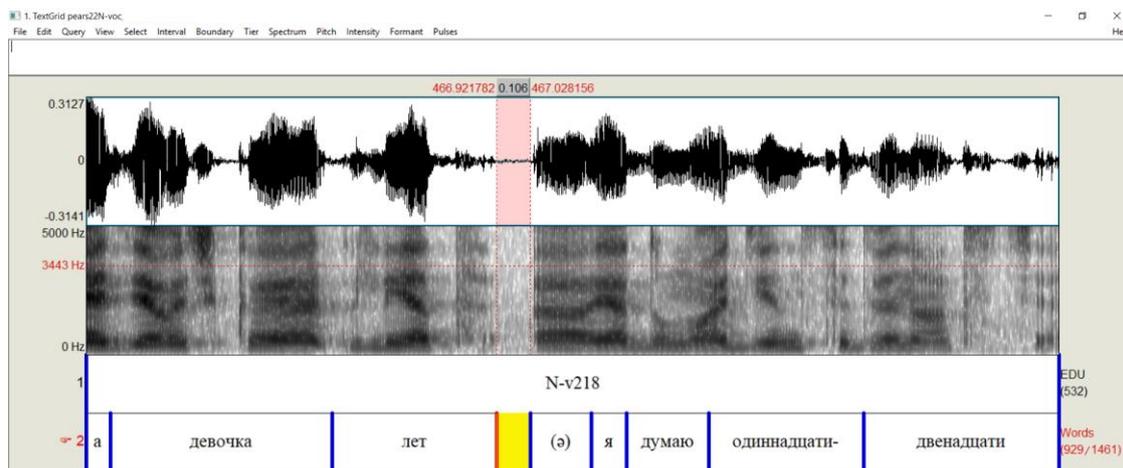


В транскриптах такие паузы записываются в отдельных строках с отдельной нумерацией (см. формат номера в примере ниже):

N-v024	/–И-и потом кладёт в такие б= /глубокие \корзины.
pN-004	()
N-v025	Сначала он собирает себе в /фартук,

Примечание. В системе нотации, разработанной для транскрипции монологической речи и описанной в Кибрик, Подлеская (ред.) 2009, такие паузы включаются во вторую по порядку ЭДЕ. Такое решение, однако, не может быть систематически реализовано в ситуации диалога.

2) При этом абсолютные паузы могут встречаться и внутри ЭДЕ:



В транскриптах они в этом случае записываются в той же строке, что и окружающий их материал:

N-v218	а /—девочка-а лет ^h () (а) я думаю /одиннадцати- \двенадцати.
--------	--

4) Паузы с шумным вдохом на границе между ЭДЕ приравниваются к абсолютным паузам (т.е. не входят ни в какие ЭДЕ), но при этом нумеруются отдельно как в текстгридах, так и в транскриптах. Все прочие заполненные паузы и разнообразные неречевые звуки, находящиеся между ЭДЕ, относятся всегда к следующей ЭДЕ, ср. пример, разобранный нами в подразделе 4.3:

N-v001	Значит /→ролик,,,
pN-002	()
N-vp001	(ц)
N-v002	(эш) () ну по →/ощущениям,
N-v003	он снят где-то-о в середине двадцатого /века,

7. Акценты, пунктуационные знаки и прочие явления, отмечаемые в транскриптах

Если явления, описанные в предыдущих разделах, отмечаются как в транскриптах, так и в текстгридах, то явления, упоминаемые в данном разделе, кодируются только в транскриптах. Подробное описание этих явлений можно найти в Кибрик, Подлеская (ред.) 2009, а также на сайте <http://www.spokencorpora.ru/>; список транскрипционных обозначений см. в разделе 8. Ниже перечисляются наиболее важные явления этой группы.

1) Акценты и движения тонов, ассоциированные с акцентами

Акцент — это выделенное произнесение словоформы в потоке речи. Дискурсивные функции акцентов состоят в первую очередь в выделении центрального элемента коммуникативной составляющей и в указании на способ связи этой составляющей с внешним контекстом (вторая функция реализуется при помощи различных движений тона). В ЭДЕ обычно выделяется главный («несущий») акцент, в наибольшей степени ответственный за встраивание ЭДЕ в локальную дискурсивную структуру.

2) Иллокутивно-фазовые значения ЭДЕ

Сюда входят указания на иллокутивную силу ЭДЕ (сообщение, вопрос, директив и проч.) и на завершенность / незавершенность высказывания. Наряду со значениями, зафиксированными в Кибрик, Подлесская (ред.) 2009, в нашем проекте отдельно размечаются случаи специального «полуутвердительного» коммуникативного намерения, при реализации которого говорящий запрашивает у других участников записи подтверждения высказываемому им предположению. Для этого значения в транскрипте используется знак “¿”.

3) Самоисправления

В транскриптах различаются ситуации, в которых самоисправление говорящего приводит к полному отказу от построения текущей ЭДЕ («сильный фальстарт»), и ситуации, в которых говорящий, несмотря на самоисправление, «спасает» текущую ЭДЕ («слабый фальстарт»).

4) Явления диалогической природы

Используемая нами система дискурсивной транскрипции была изначально разработана на материале монологической устной речи. В целом она применима и для разметки диалогического текста, однако определенные ее элементы требуют уточнения. Выше, в подразделе 6.3, уже обсуждалось одно важное отличие, связанное с обозначением пограничных абсолютных пауз. См. также рассмотренное выше в пункте 2 обозначения для иллокуции «полуутверждения». В транскриптах диалогического дискурса также необходимо размечать следующие явления, не характерные для монолога:

- обрыв ЭДЕ говорящего, вызванный не самоисправлением, а потерей очереди хода (говорящего «перебивают»);
- совместное построение реплик (один говорящий начинает реплику, другой говорящий ее заканчивает).

Для первого явления используется обозначение “≈”, для второго — “%” в конце / начале ЭДЕ.

Примечание. Еще одно важное диалогическое явление — это наложения реплик. В транскриптах, описываемых в настоящей инструкции, наложения специально не отмечаются: они «вычисляются» на основании информации, содержащейся в текстгридах. См. представленные на сайте примеры партитурных транскриптов.

5) Цитации, парентезы

6) Ускоренный темп произнесения, сниженный регистр, эмфатическое произнесение

7) Редуцированное произнесение

8) Удлиненная реализация фонем

9) Дополнительные особенности произнесения (смычки, придыхания и проч.)

10) Комментарии

Наблюдения транскрайберов, касающиеся явлений, не отмечаемых систематически в транскриптах, указываются в отдельном столбце таблицы.

8. Список транскрипционных обозначений

Ниже приведен список транскрипционных обозначений, используемых при вербально-просодической разметке. Отметим, что в список не включены дополнительные обозначения, применяемые при отображении разметки в формате окончательных раздельных, совместных и партитурных транскриптов.

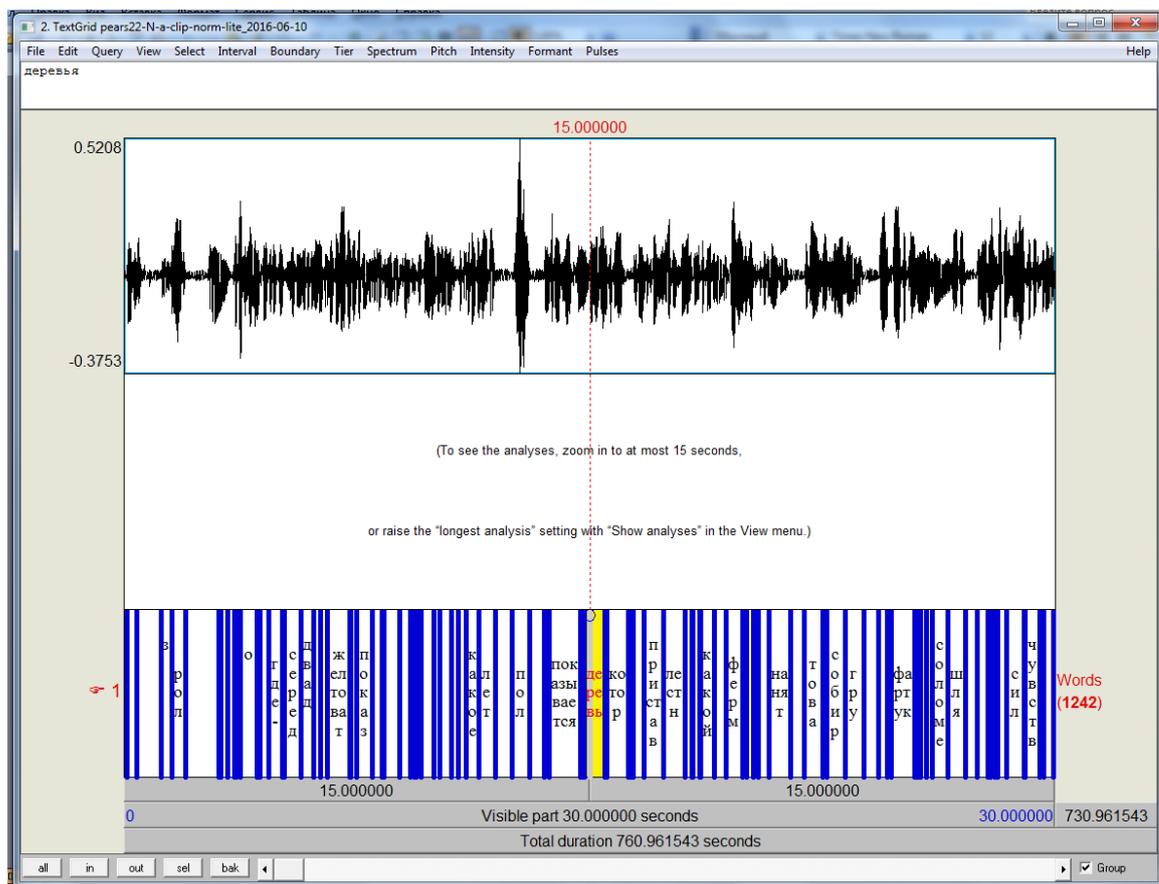
Обозначение	Явление
Деление на строки	Деление на элементарные дискурсивные единицы (ЭДЕ)
()	Абсолютные паузы
(ə), (v), (ш), (?), (эш) и проч.	Заполненные паузы различной природы
{sp}, {cl}, {cg}, {laugh}	Неречевые звуки говорящего
/ \ — ^ и т.д. (перед словом)	Наличие акцента на словоформе и характер движения тона на ударном слоге акцентированной словоформы
↑ ↓ →	Коммуникативно значимое движение тона вне ударного слога акцентированной словоформы
Подчеркивание ударной гласной акцентированного слова	Несущий (главный) акцент ЭДЕ
Заглавная буква в начале строки	Начало нового предложения
.	Завершение иллокуции сообщения
?	Завершение иллокуции вопроса
!	Завершение иллокуции директива
@	Завершение иллокуции обращения
!	Завершение иллокуции «полуутверждения» (запроса на подтверждение предположения)
!	Восклицательность
...	Значение неполноты информации, совмещенное с завершением иллокуции
,	Стандартная незавершенность иллокуции
””	Значение неполноты информации, совмещенное с иллокутивной незавершенностью
:	Незавершенность, восполняемая последующим контекстом
« »	Начало и конец прямой цитации
—	Начало и конец сплита (разрыва ЭДЕ)
()	Начало и конец парентезы (вставки предложения внутри другого предложения)
(*	«Односторонняя парентеза» (начатая, но не завершенная вставка предложения внутри другого предложения)
=	Обрыв слова
	«Слабый фальстарт» (точка прерывания при самоисправлении внутри ЭДЕ)
==	«Сильный фальстарт» (точка прерывания при самоисправлении на границе ЭДЕ)
~	Обрыв ЭДЕ, которая расценивается говорящим как уместная и не подлежащая исправлению (апосиопеза)
≈	Обрыв ЭДЕ, вызванный потерей хода говорящим
%	ЭДЕ данного говорящего является частью совместно построенной реплики
ʔ	Гортанная смычка
ʋ	Губная смычка
h	Придыхание

°	Гласный призвук в начале или конце слова
а-а с-с й-я я-а	Удлиненная реализация фонем
Полужирный шрифт	Эмфатическое выделение
Выделение <i>серым цветом</i>	Редуцированное произнесение
<i>Курсив</i>	Ускоренный темп
Уменьшенный кегль	Сниженный тональный регистр
" "	Названия фильмов, машин и т.д.
< >	Предположительная транскрипция неразборчивого фрагмента
< >	Варианты транскрипции неразборчивого фрагмента
<НРЗБ2>	Неразборчивый фрагмент с указанием числа слогов
> <	Разборчивый, но неидентифицируемый фрагмент
>>	Технический обрыв в начале фрагмента или рассказа
<<	Технический обрыв в конце фрагмента или рассказа

9. Ликбез по работе с аннотациями в Praat

В этом разделе описаны общие принципы работы с аннотациями (текстгридами) в Praat.

- 1) Запускаем программу Praat (программу можно скачать по адресу <http://www.fon.hum.uva.nl/praat/>, после этого крайне желательно почитать хелп и самостоятельно разобраться с основными доступными функциями; ниже будет описание некоторых стандартных действий, но все же оно не может полностью заменить хелп; ниже описываются правила и приемы работы под Windows, но все то же самое можно делать и на Mac – с незначительными и чаще всего очевидными отличиями).
- 2) Загружаем в меню объектов (меню Open > Read from File или, если файл слишком большой, Open > Open long sound file; нужно ли вам пользоваться этой опцией, зависит от размера оперативной памяти вашего компьютера) требуемый звуковой файл (формат .wav)
- 3) Загружаем в меню объектов (через меню Read from file) требуемый текстгрид.
- 4) Выделяем оба файла (аудиофайл и текстгрид) в меню объектов и нажимаем кнопку View & Edit в правой панели. Вид появившегося окна зависит от наших настроек. Если мы нажимаем на эту кнопку впервые на данном компьютере, то должно появиться примерно такое окно:



В этом окне отображаются:

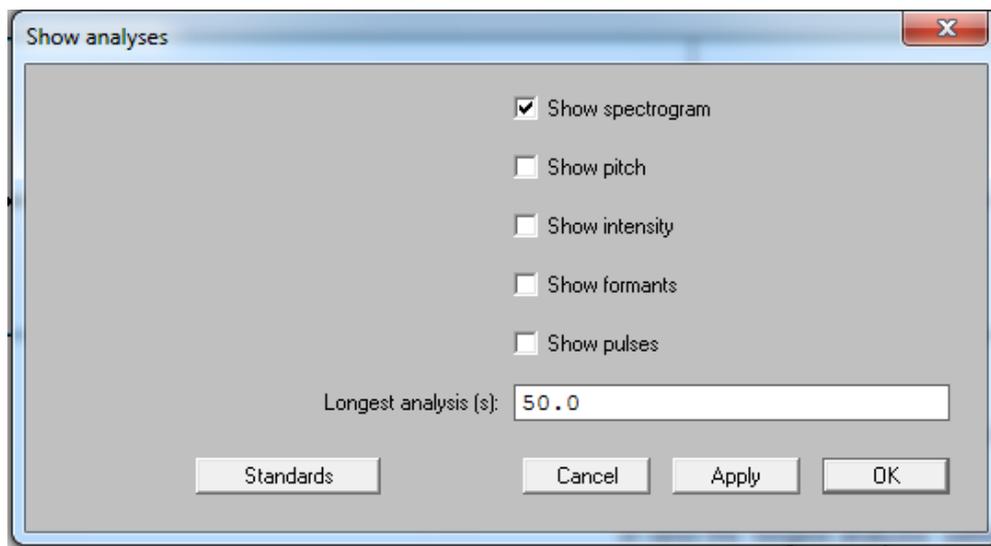
- осциллограмма
- область анализа (спектрограмма, тональная кривая и др.)
- слои аннотации (на скриншоте показан один слой — Words)

По умолчанию в области анализа визуальное представление строится, только если в окне просмотра отображается 10 секунд звучания или менее.

Прежде чем приступать к непосредственной работе, нужно изменить некоторые настройки.

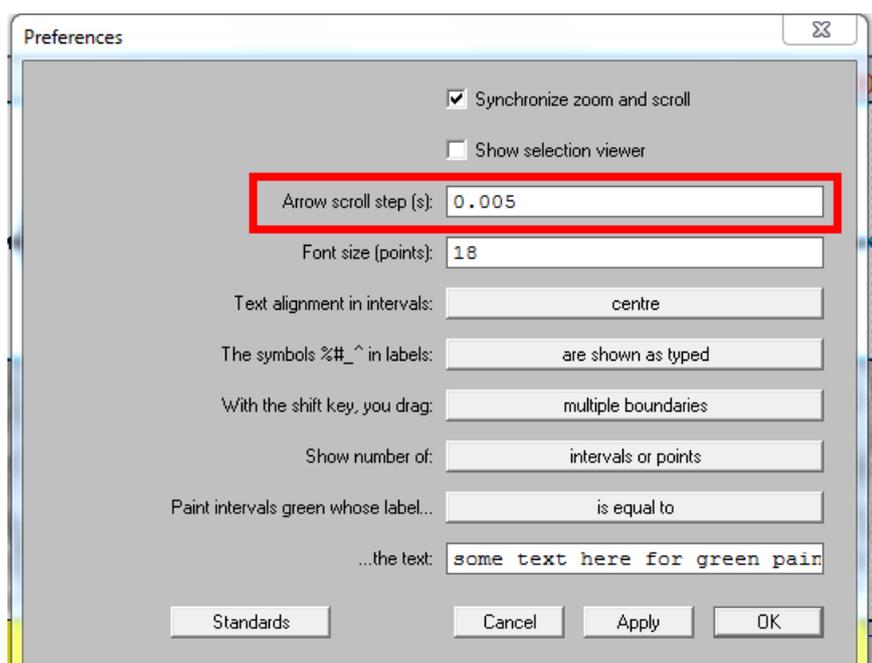
А. Меню View > Show analyses...

Включить галочку Show spectrogram, снять все другие галочки, в поле Longest analysis (s) указать значение побольше (я у себя делаю 50 секунд).



Б. Меню File > Preferences

Необходимо изменить значение параметра Arrow scroll step (s), оно должно равняться 0.005 (по умолчанию стоит существенно более высокое значение, это для наших задач крайне неудобно).



5) При открытии звукового файла (неважно, есть у него textgrid или нет) по умолчанию всегда видно не более 30 секунд. Для изменения размера / наполнения видимой области окна можно использовать следующие кнопки / клавиши:

- Кнопка **all** в нижнем левом углу – при ее нажатии в окне отображается весь звуковой файл (при этом область «анализов» может стать недоступной, если продолжительность файла превышает установленный лимит)
- Кнопка **in** в нижнем левом углу – увеличение масштаба: фрагмент файла, отображаемый в окне, становится в два раза менее длительным; при этом «отсекается» одинаковая длительность слева и справа (и это иногда может быть не слишком удобным)
- Кнопка **out** в нижнем левом углу – уменьшение масштаба, т.е. функция, обратная **in**

- Кнопка **sel** в нижнем левом углу – работает при наличии выделенного интервала в осциллограмме или «анализах» (как выделить интервал – см. ниже); при ее нажатии ранее выделенная область занимает все окно. В принципе, это **самая важная** нам кнопка: с ее помощью можно наиболее тонко контролировать видимую область файла.

- Кнопка **bak** в нижнем левом углу – возврат к предыдущему виду окна

- **Ползунок** справа от кнопок (или клавиши PgUp / PgDn) – «перемотка» файла без изменения размера видимой области. Разумеется, это тоже совершенно необходимая для работы функция.

б) Работа с курсором

- Курсор можно установить в любом месте видимого окна (в области осциллограммы или «анализов») мышью, далее его можно передвигать стрелками вверх/вниз на клавиатуре (шаг, на который курсор при этом сдвигается, мы установили в меню Preferences)

- Чтобы выделить интервал, можно воспользоваться Drag&Drop или, установив курсор в одном месте, затем нажать Shift и, не отпуская, его установить вторую границу интервала

- При наличии выделенного интервала двигать его границы можно при помощи следующих клавиш / сочетаний клавиш:

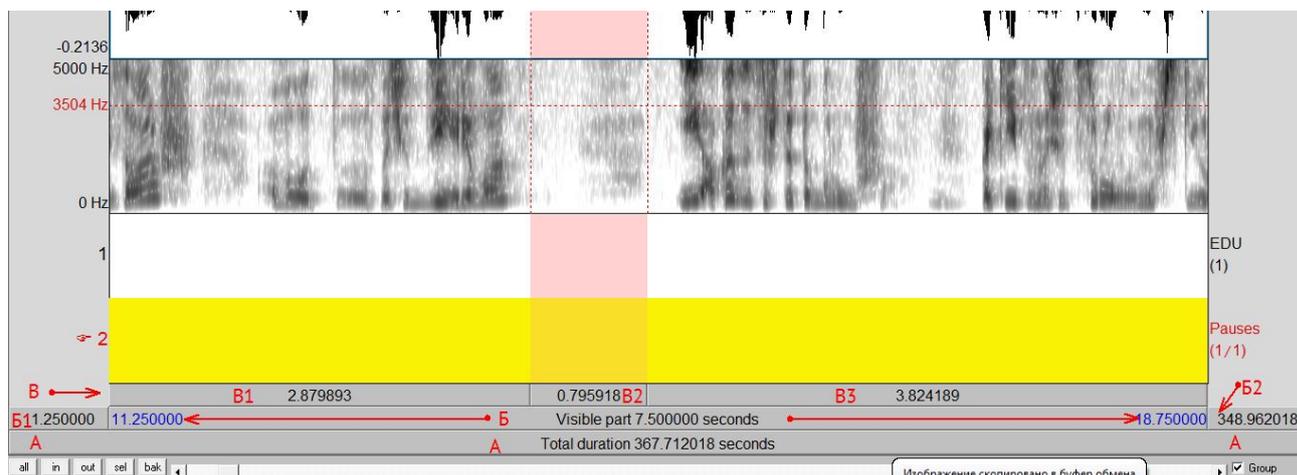
-- Клавиши «вниз / вверх» перемещают соответственно вправо / влево одновременно и левую, и правую границу интервала.

-- Сочетание Shift + «вниз / вверх» перемещает соответственно вправо / влево только левую границу интервала.

-- Сочетание Ctrl + «вниз / вверх» перемещает соответственно вправо / влево только правую границу интервала.

7) Проигрывание нужных фрагментов файла

Непосредственно под всеми представлениями, относящимися к звуковому файлу, имеются три кликабельные полосы – см. рисунок ниже:



- В нижней полосе (**А**) указана информация об общей продолжительности звукового файла. Нажатие на любое место на этой полосе запускает проигрывание всего файла с начала (чтобы остановить проигрывание любого фрагмента, нужно нажать Esc).

- В средней полосе (**Б**) указана информация о длительности, а также времени начала и конца текущего окна; нажатие на область **Б** запускает проигрывание фрагмента звукового файла от начала до конца текущего окна.

-- В области **Б1** (слева от **Б**) указана длительность от начала файла до начала текущего окна; нажатие на **Б1** запускает проигрывание этого (невидимого) фрагмента файла.

-- В области **Б2** (справа от **Б**) указана длительность от конца текущего окна до конца файла; нажатие на **Б2** запускает проигрывание этого (невидимого) фрагмента файла.

- В верхней полосе (**В**) указывается информация о длительности отрезков:

-- от начала видимого окна до левого курсора (**В1**)

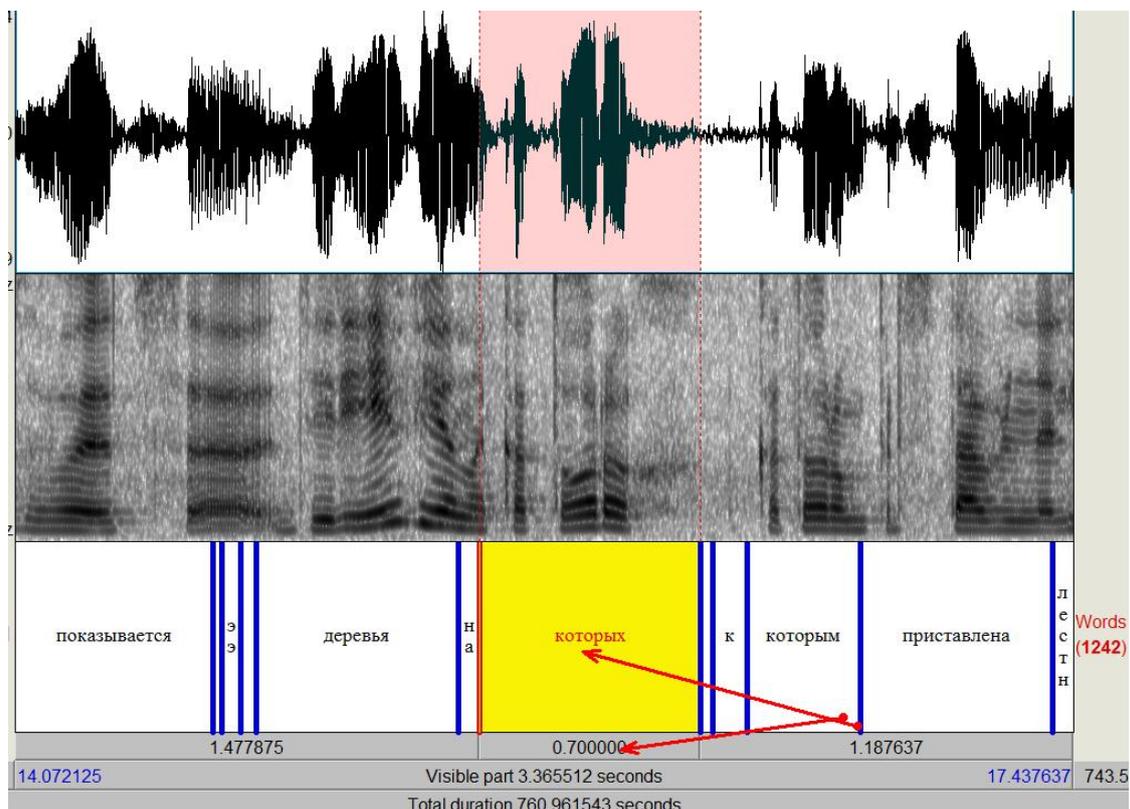
-- от левого до правого курсора (**В2**)

-- от правого курсора до конца видимого окна (**В3**)

Разумеется, нажатие на эти области также запускает проигрывание соответствующих фрагментов файла.

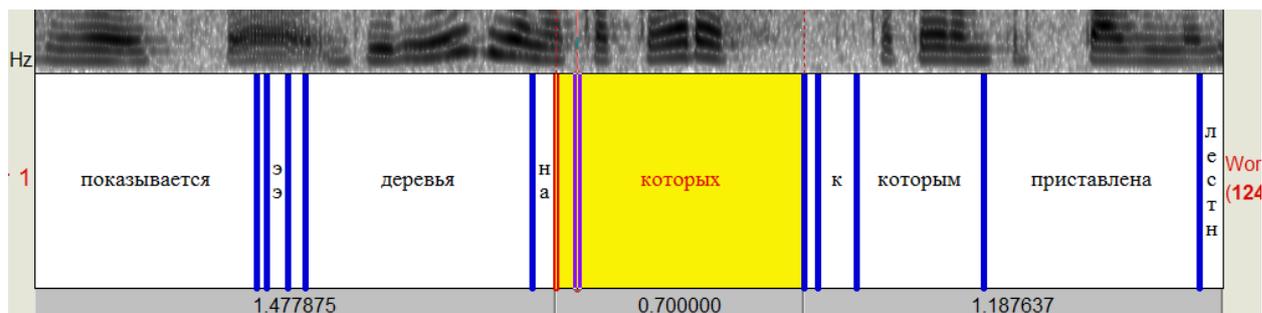
8) Проигрывание интервала в нужном слое

Если нажать на любой интервал в любом слое аннотации, то автоматически выделится область, соответствующая этому интервалу. И ее можно проиграть стандартным способом:



9) Изменение границ интервала

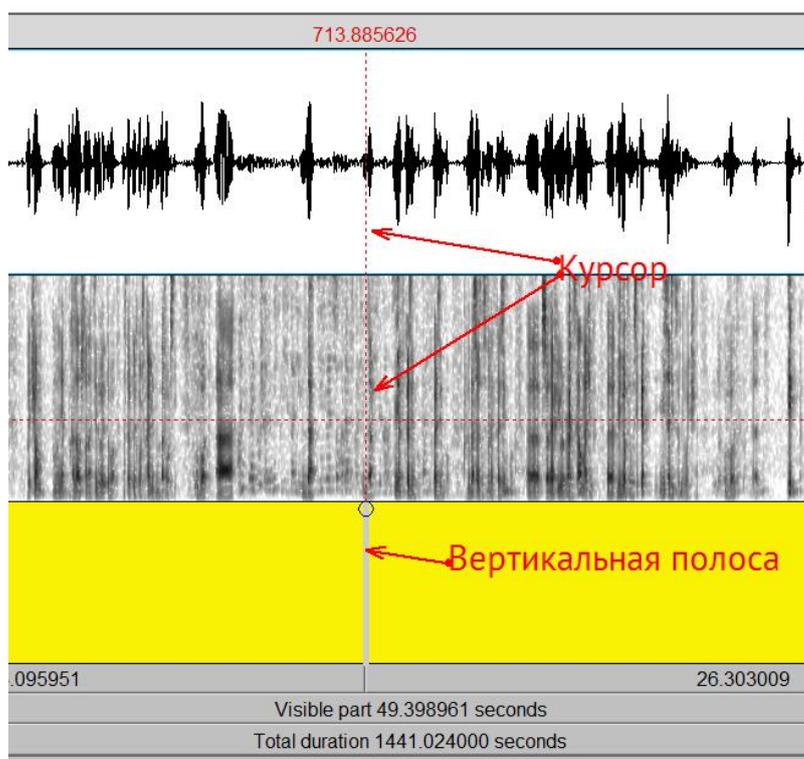
Чтобы передвинуть границу интервала, достаточно нажать на границу и переместить ее на новое место при помощи Drag&Drop.



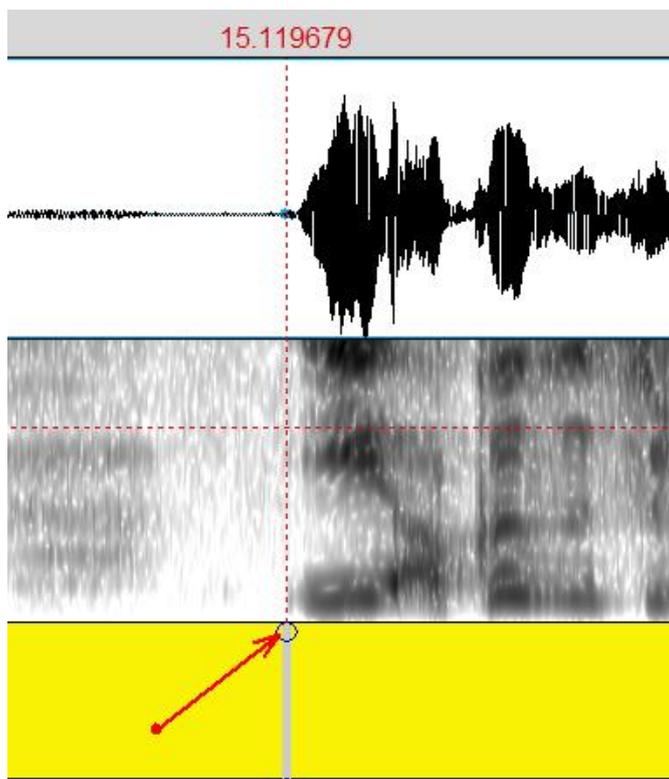
Правда, это не самый надежный опыт. Иногда полезнее поставить новую границу в каком-нибудь другом слое, а затем скопировать эту границу в целевой слой.

10) Создание новой границы

Чтобы создать границу, достаточно поставить курсор на нужное место в осциллограмме или спектрограмме. В слое аннотации появится вертикальная серая полоса:



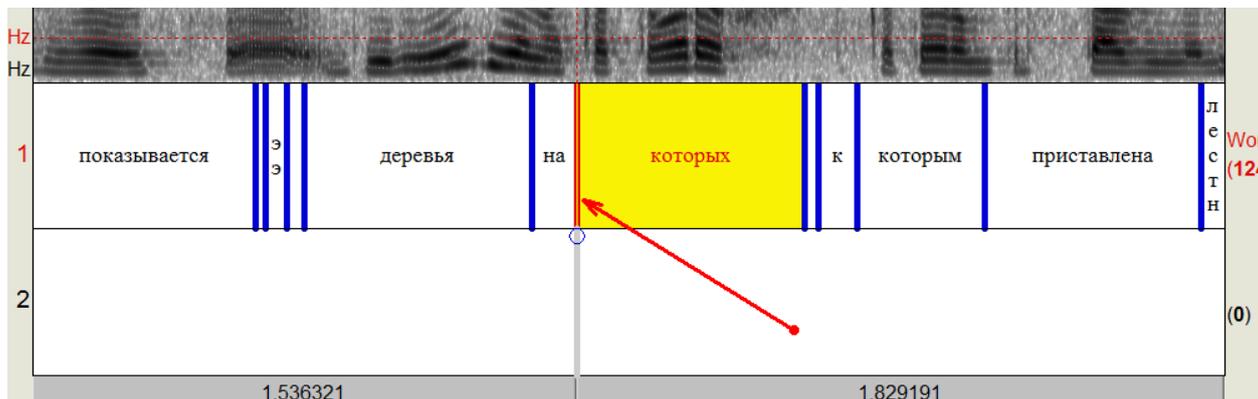
При перемещении курсора эта полоса также будет перемещаться. Когда положение курсора вас удовлетворяет, нужно нажать на серый кружок вверху вертикальной линии, продолжающей курсор в уровне аннотации (**вместо нажатия на серый кружок также можно просто нажать на Enter**):



Линия после этого станет красной.

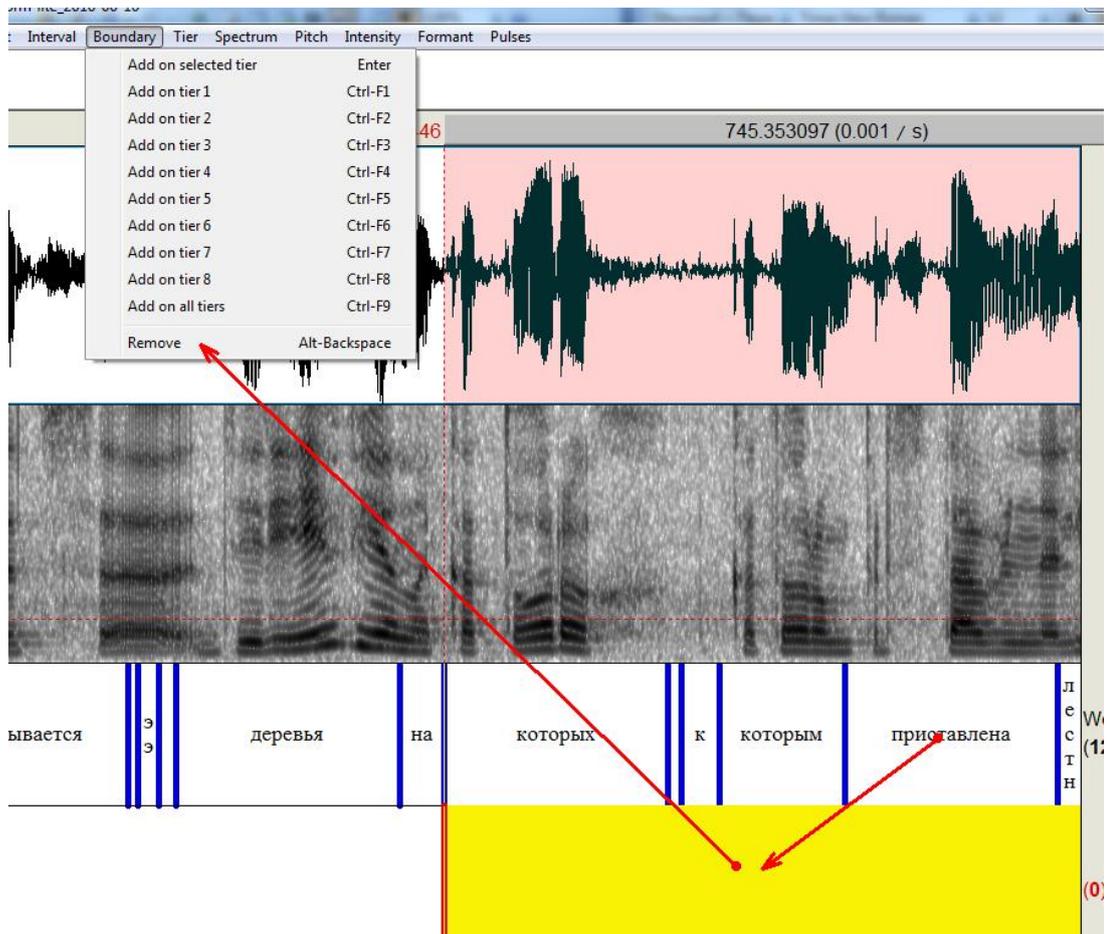
11) Копирование границы из одного слоя в другой

Чтобы скопировать границу из одного слоя в другой (а это нужно будет делать при заполнении нового слоя EDU — об этом ниже), достаточно нажать на границу в слое-источнике, после чего во всех остальных слоях появится уже знакомая серая полоса с кружочком. При нажатии на кружочек в нужном слое появится граница, полностью совпадающая по времени с границей из слоя-источника.



12) Удаление границы

Чтобы удалить границу, нужно нажать на интервал справа от границы и после этого воспользоваться меню Boundary – Remove. Также можно просто нажать Alt-Пробел.



13) Создание нового слоя аннотации

Чтобы создать новый слой аннотации, находясь в режиме просмотра и редактирования текстгрида, необходимо войти в меню Tier > Add interval tier... Далее в поле Position указать порядковый номер слоя, в поле Name — ввести название слоя.