

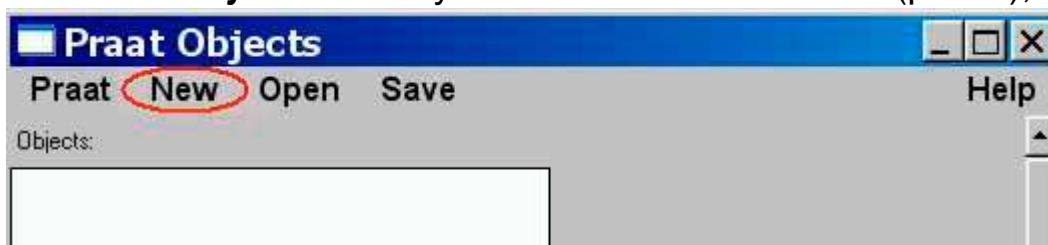
Рекомендации
по применению программы PROSOGRAM
(Mertens 2004)

В принципе, программа может работать только с собственно акустическим сигналом, но в этом случае она формирует результат для «вокальных ядер», определяемых по внутренним правилам программы, и в категорию гласных могут попасть и сонанты тоже.

Поскольку в нашей концепции коммуникативно значимые тональные изменения можно наблюдать только на гласных, помимо файла с самим речевым сигналом программе надо дать (поместить в одну с ним директорию) файл <имя речевого файла>.textgrid, в котором будут представлены временные координаты гласных в анализируемом сигнале, т.е. разметка интервалов гласных в данном сигнале. Файл этот должен быть в формате Praat. Поэтому удобнее всего создавать его средствами именно этой программы. Координаты гласных и их фонетические маркеры можно получить и в какой-нибудь другой программе, но потом результаты придется переформатировать в формат Praat, либо, работая одновременно с двумя программами, устанавливать границы сегментов в Praat на основании координат в другой программе (например, в WinSnoori или SpeechAnalyzer).

Чтобы сразу получить этот файл в требуемом формате, следует действовать следующим образом.

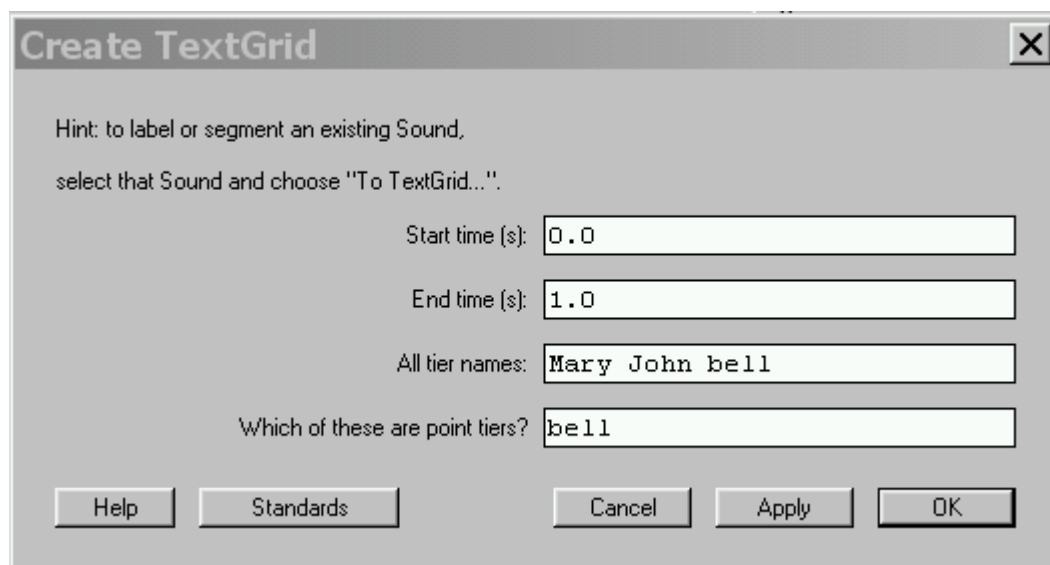
- войти в программу Praat;
- в окне Praat **Objects** “кликнуть” мышкой на поле **New** (рис. 1);



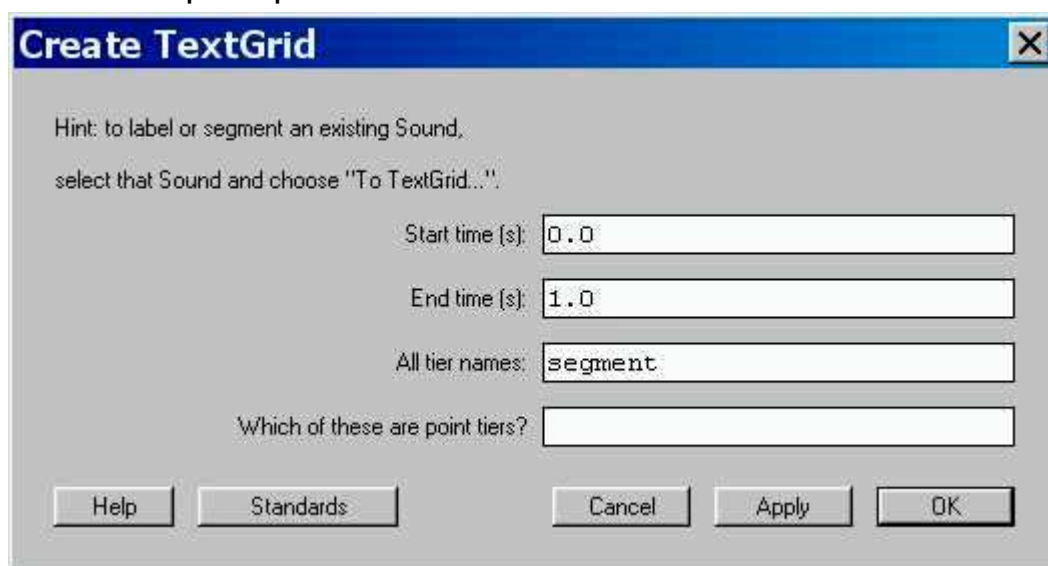
- в открывшемся меню «кликнуть» мышкой на поле **Create TextGrid**;



– заполнить необходимые поля в окне **Create TextGrid**:

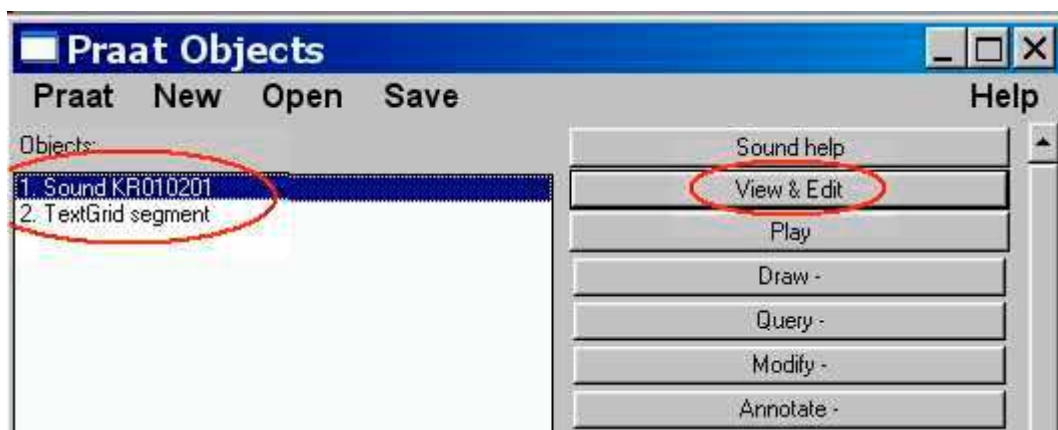


в поле **End time** указать длительность, несколько превышающую длительность обрабатываемого сигнала; остальные поля должны быть заполнены как в примере ниже

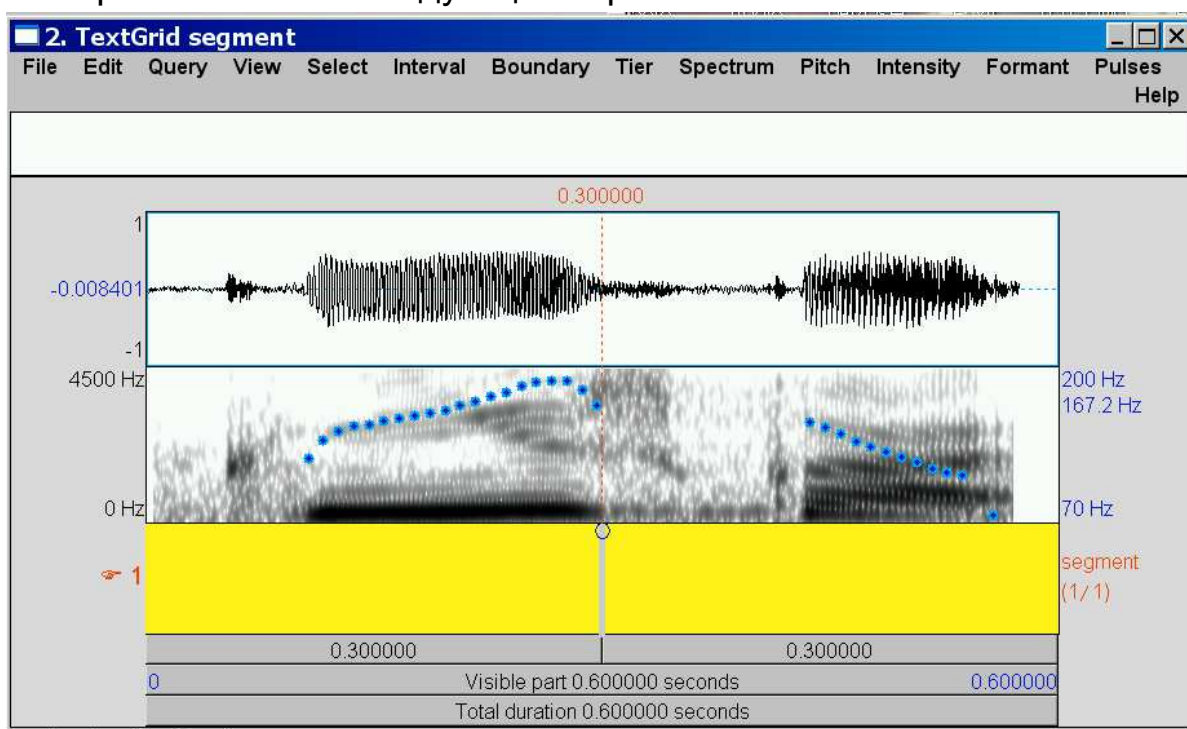


– “сказать OK”

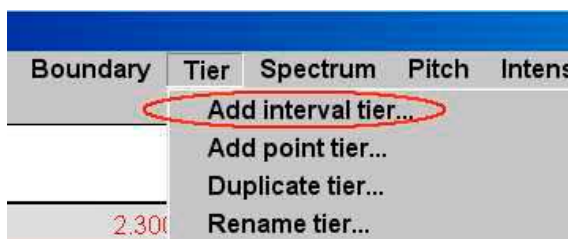
Теперь в окне Praat **Objects** открыть меню **Open** и ввести имя файла с обрабатываемым речевым сигналом. После этого выделить строки с файлом сигнала и **TextGrid**-файлом и запустить команду **“View&Edit”**.

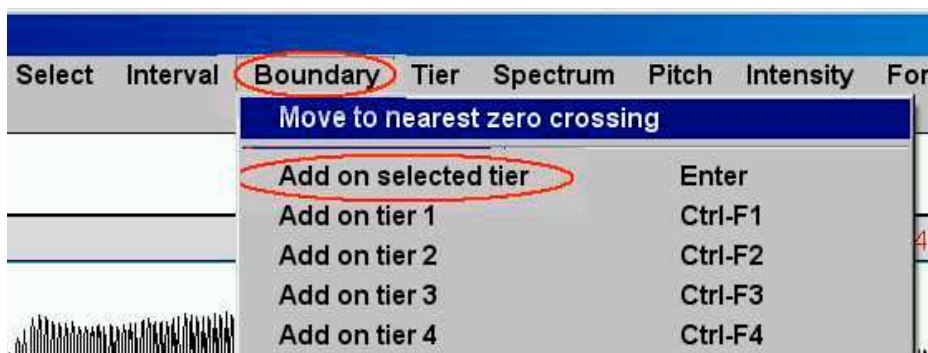


На экране появится следующая картинка.



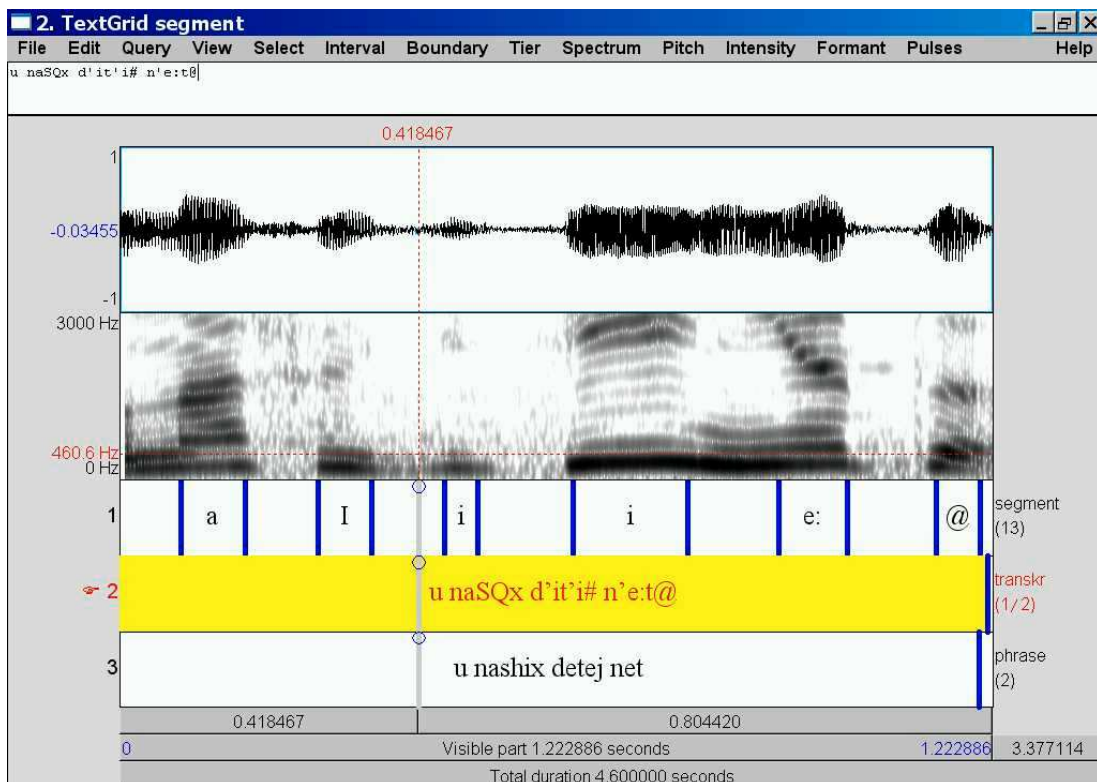
Теперь необходимо добавить еще два уровня (tier) описания обрабатываемого речевого отрезка: уровень транскрипционного и орфографического (в латинской транслитерации русского текста) описания. Для этого, выбрав пункт меню **“Tier”**, следует последовательно дважды выполнить команду **“Add interval tier”** (см. рис.) и задать для уровня 2 имя **“transcr”**, а для уровня 3 – **“phrase”**.





Теперь надо последовательно выделять один межпаузный интервал за другим и для каждого выполнять команды **Boundary** и **Add on selected tier**, вставляя для уровня “**transcr**” транскрипцию данного интервала (ее следует брать из файлов <имя речевого сигнала>.uns), а для уровня “**phrase**” – соответствующий русский текст в латинской транслитерации (для работы в Praat можно вставлять текст и в русской орфографии, но PROSOGRAM кириллицы не понимает, и при окончательном оформлении графиков придется кириллический текст вставлять средствами графических редакторов).

На последнем этапе последовательно выделяется один гласный за другим, для каждого выполняются команды **Boundary** и **Add on selected tier** и соответствующие участки **tier “segment”** маркируются принятыми в конкретном исследовании символами гласных (они уже есть в **tier “transcr”**, кроме «ы», для которого на этом уровне используется заглавное “I” (I) вместо принятого нами “Q”). Должна получиться следующая картинка.



На всех этапах надо постоянно сохранять результаты заполнения этого текстового файла. Для этого выбирать меню "**File**", а в нем выполнять команду "**Save TextGrid as a text file**". При этом имя этого текстового файла должно совпадать с именем файла с обрабатываемым сигналом, а расширение должно быть **TextGrid**.

Теперь все готово, чтобы вызывать программу PROSOGRAM. Делать это надо в соответствии с описанием этой программы.

Все файлы скриптов PROSOGRAM уже должны быть размещены на компьютере в рабочей директории программы (например, c:\prosogram\ или d:\prosogram\ или в любой другой директории по выбору исследователя).

Кроме того, на компьютере должны быть установлены программы *Ghostscript* и *Gsview*, причем обе должны быть установлены в директорию (папку) с именем GS, а в скрипте **prosomain.praat** должна быть строка (например,

```
path_ghostscript$ = "d:\Program Files\gs\gs9.10\bin\gswin32c"
```

с указанием места расположения соответствующей версии программы *Ghostscript*.

Прежде чем вызывать PROSOGRAM в окне Praat **Objects** должны быть загружены три файла:

<имя_обрабатываемого_сигнала>.wav

<имя_обрабатываемого_сигнала>.TextGrid

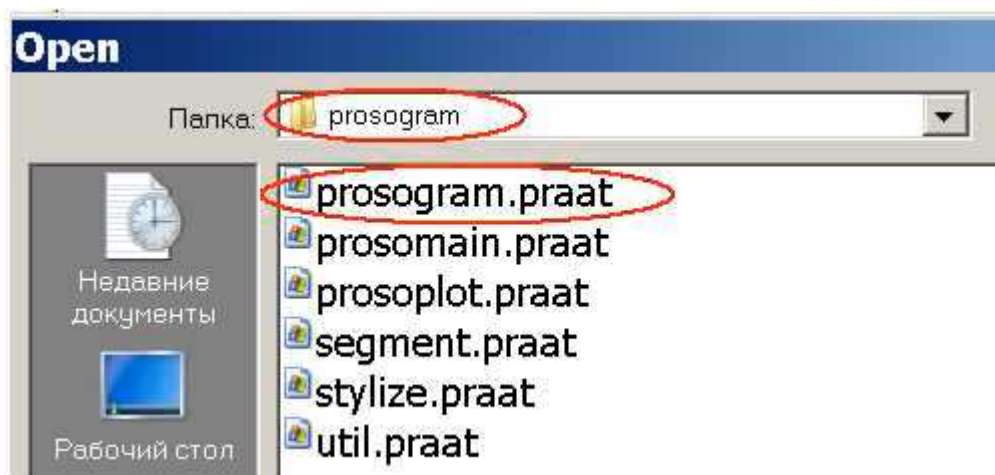
<имя_обрабатываемого_сигнала>.pitch



Теперь в окне Praat **Objects** открыть меню Praat, а в нем подменю **Open Praat script**.



В открывшемся окне найти на дисках папку, в которой хранится PROSOGRAM, и загрузить скрипт ***prosogram.praat***.



По умолчанию prosogram.praat использует на графиках шрифт Helvetica. Если при транскрибировании в латинице нужно, чтобы различались заглавное «и» (I) и строчное «л» (l), предпочтительнее использовать гарнитуру TimesNewRoman. Для этого надо отредактировать скрипт: исходный текст

```
# Font used in plot
font_family$ = "Helvetica" ; else use "Times"
заменить следующим
# Font used in plot
# font_family$ = "Helvetica" ; else use "Times"
font_family$ = "Times" ; else use "Helvetica"
```

После вызова скрипта следует прежде всего настроить параметры обработки.

В разделе ***Segmentation method*** надо задать ***Using external segmentation in tier "segm..."***, поскольку мы собираемся обрабатывать заранее нами выделенные интервалы гласных.

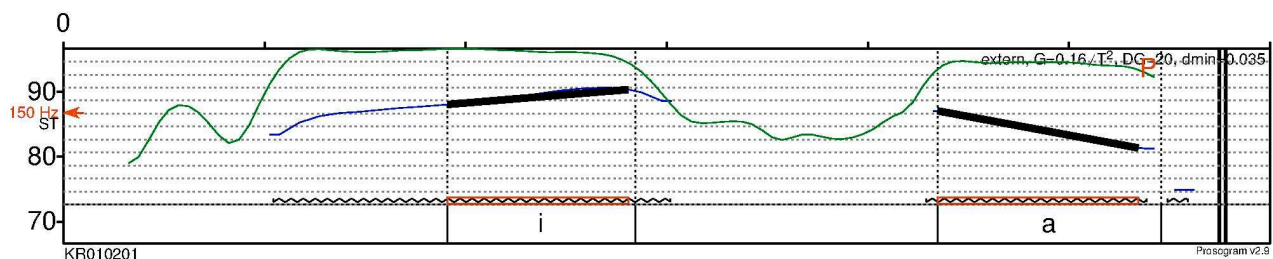
В разделе ***Thresholds*** выбираем $G=0.16/T^2$, $DG=20$, $dmin=0.035$. Это начальный вариант. Потом можно будет подобрать наиболее соответствующий обрабатываемому материалу и задачам исследования вариант.

В разделе ***Time interval per strip*** лучше всего указывать слегка увеличенную длительность обрабатываемого сигнала в секундах (при этом нельзя опускать «ведущий» ноль, т.е. надо задавать 0.6, а не .6).

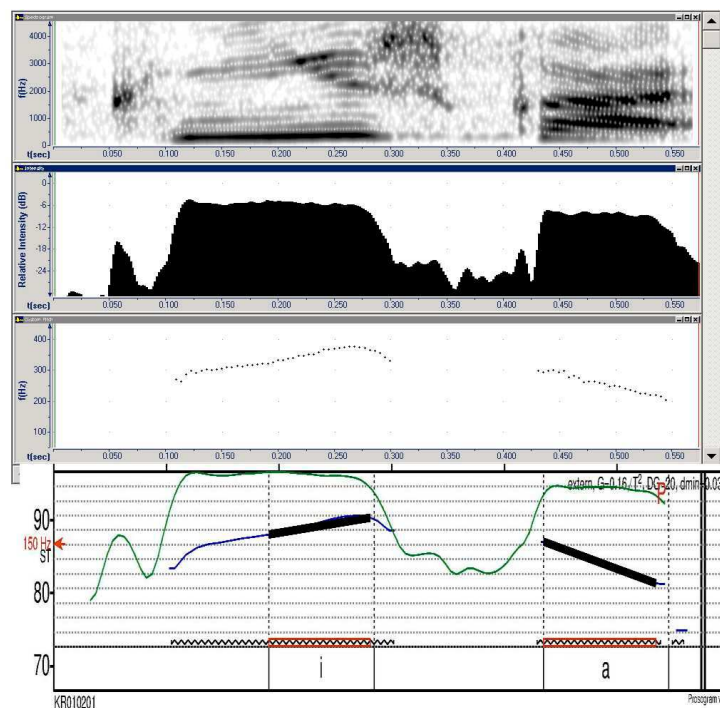
Output format лучше всего задавать в виде ***EPS and JPG 300 dpi***, т.к. в этом случае сразу получается графический файл в jpg-формате, что позволит потом просто совмещать полученную просограмму со спектрограммой, взятой из другой программы.

Но если желательно на просограмме сохранить текст в кириллице уже при ее создании, а не вставлять его потом с помощью графических редакторов, то выбрать следует **EMF (Windows Enhanced Metafile)**, поскольку используемая для создания jpg-файла программа не понимает кириллицы. Получившийся emf-файл с просограммой можно спокойно вставлять в тексты, создаваемые Word'ом, а для преобразования в jpg-файл можно использовать редактор изображений XnView: его можно скачать из Интернета (даже русифицированную версию) и использовать бесплатно для научных и учебных задач. Одновременно с преобразованием желательно вдвое уменьшить размер картинки.

По окончании работы программы получится вот такая картинка.



А после совмещения с результатами анализа программой SpeechAnalyzer спектральных характеристик и огибающей интенсивности получится следующее:



Для начала надо получить рисунок с результатами анализа этого же сигнала в SpeechAnalyzer'е, которые должны быть расположены на

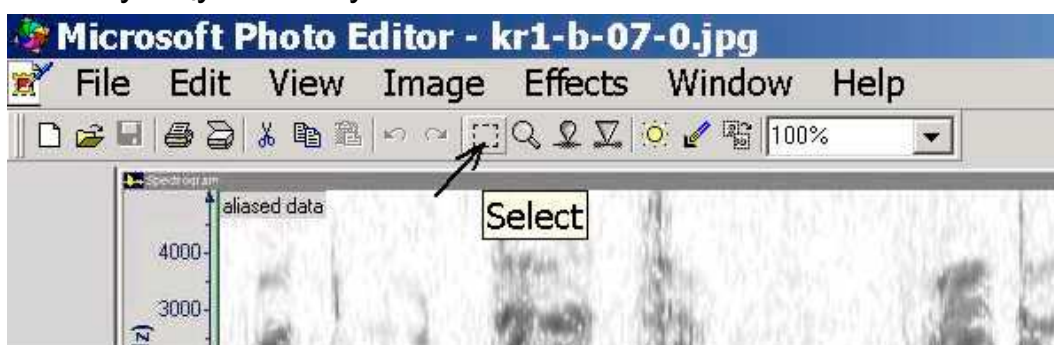
рисунке в следующем порядке (сверху вниз): динамическая спектрограмма, контур интенсивности, мелодический контур.

Для совмещения оба графика должны быть приведены к единому масштабу по оси времени (в пикселях на секунду). Поскольку программы выдают графики в разном масштабах, приходится деформировать одну из картинок, изменяя ее размер по горизонтали (размер по вертикали остается неизменным).

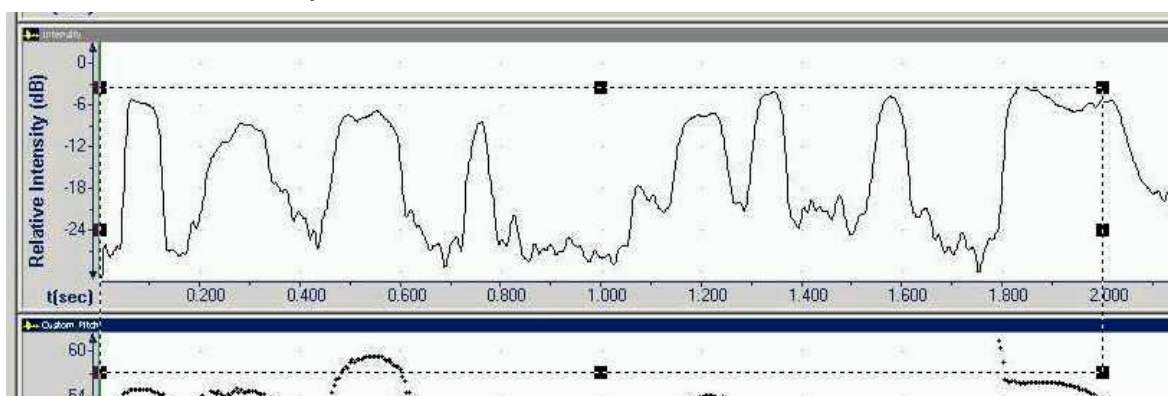
Делать это можно в любом графическом редакторе, но удобнее всего – в Microsoft photo editor.

Для этого надо загрузить в него нужные картинки и измерить в пикселях расстояние между точками на оси времени: от нуля секунд, до одной, двух, трех секунд. Удобнее всего делать это следующим образом.

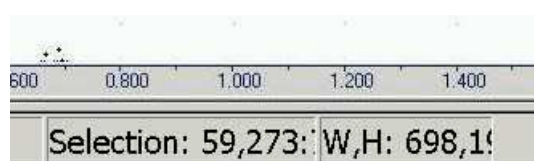
Запускается режим «выделения», «кликнув» мышкой на соответствующую иконку.



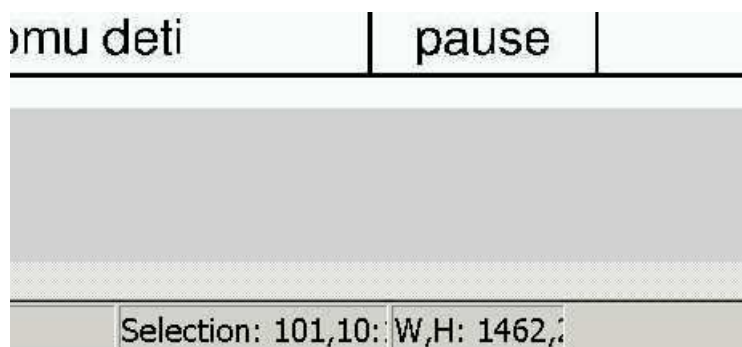
Движением мышки при нажатой левой клавише задается поле выделения, и его левая и правая границы совмещаются с выбранными точками на шкале времени.



Ширина поля выделения прочитывается в окне в середине нижнего края рисунка (W,H: xxx).

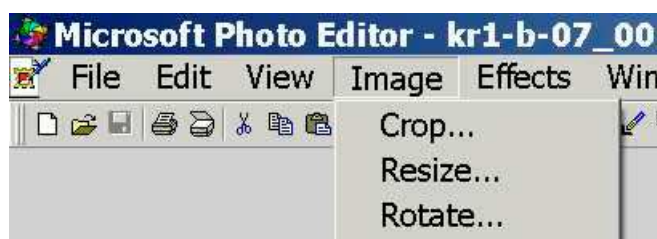


Такая же операция проделывается с просограммой.



Вычисляется коэффициент «сжатия» для просограммы или «расширения» для спектрограммы в зависимости от предпочтений исследователя. Иногда лучше сжимать просограмму, т.к. при расширении спектрограммы становится менее привлекательным ее вид.

Для этого надо открыть меню "Image" и выполнить команду "Resize".



В открывшейся панели надо проставить «галочки» как указано на рисунке, а в окошке "Width" установить вычисленный коэффициент деформации (целым числом в процентах), после чего «нажать» ОК.



Теперь создать новый рисунок (файл) размером 2000x1500 пикселей и последовательно скопировать в него рисунок со спектрограммой и просограмму (один из них уже измененного размера), с максимально возможной точностью совместив отсчеты на соответствующих осях времени. Потом из получившегося рисунка с небольшими белыми полями выкопировать только нужную часть и сохранить ее в виде графического файла с требуемым именем.

Можно повысить точность совмещения шкал времени путем указания в поле Width не величины коэффициента сжатия (расширения)

в процентах, а требуемой новой ширины в пикселях. Для этого вычисленный коэффициент сжатия (расширения) со всеми знаками после запятой умножается на исходную ширину (в пикселях) изменяемого рисунка и округленный до целой величины результат записывается в поле Width.

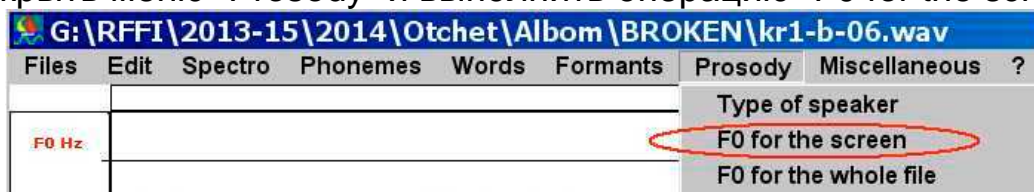
В описанном выше варианте работы мелодический контур соответствует результатам измерения частоты основного тона самой программой PROSOGRAM. Существует возможность строить мелодический контур по результатам измерения частоты основного тона программами Praat, SpeechAnalyzer или WinSnoori.

NB!! WinSnoori и SpeechAnalyzer не “умеют” измерять частоту О.Т. выше 500 Гц.

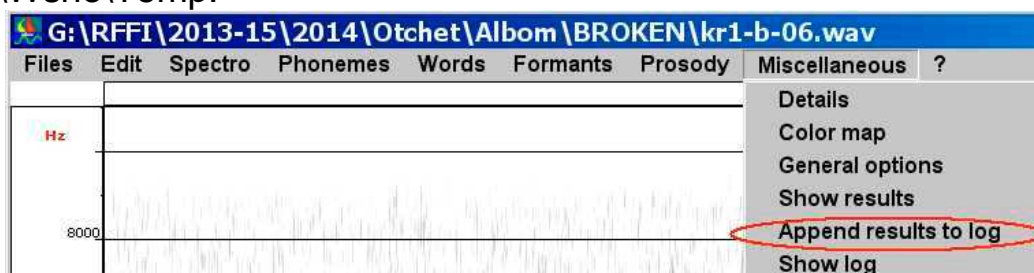
Если контур О.Т. определяется программой Praat, то его следует сохранить в файле с именем <имя_клаузы>.pitch и затем использовать при создании просограммы.

При использовании программы WinSnoori поступают следующим образом:

- загружается анализируемый файл в программу;
- открыть меню “Prosody” и выполнить операцию “F0 for the screen”.



- сохранить полученные результаты в текстовом файле; для этого открыть меню “Miscellaneous” и выполнить операцию “Append results to log”; результат будет записан в файл log.txt в папке \Program files\Wsno\Temp.

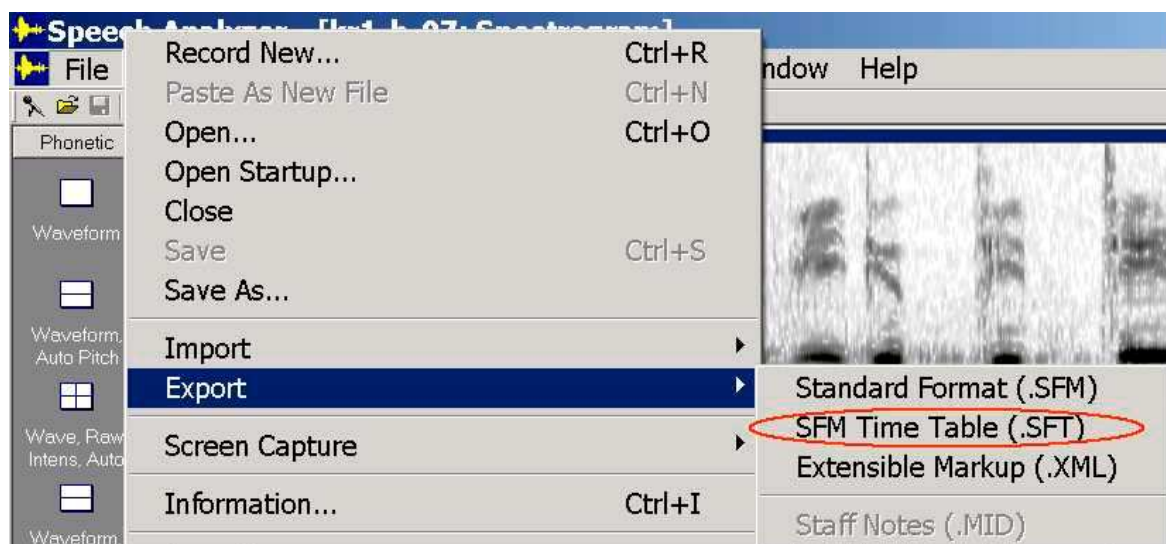


- этот файл надо перенести (не скопировать, а перенести, чтобы он не оставался в этой папке) в папку, где находится звуковой файл анализируемой клаузы, изменив его имя на «имя_клаузы».win;
- теперь win-файл (с именем не длиннее 8-и символов) надо превратить с помощью программы addpros.exe в новый файл «имя клаузы».wsn; после чего перенести его в папку, где находится анализируемый звуковой файл, заменив при этом расширение wsn на pitch.

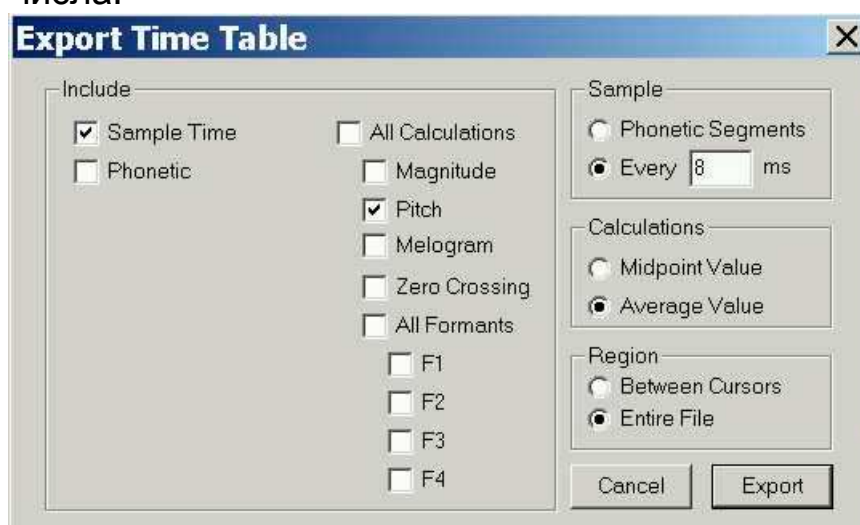
NB!! Программа ADDPROS.exe создана в операционной системе DOS. В системе Windows запускать ее следует из командной строки, соблюдая при этом следующее правило: имя

обрабатываемого файла не должно быть длиннее 8-и символов, а расширение – трех.

При желании (или при необходимости), можно использовать данные о частоте основного тона, которые дает программа SpeechAnalyzer (используются только данные smooth pitch). Для этого в меню “File” надо открыть подменю “Export” и выполнить команду “SFM Time Table (.sft)”.



В открывшейся панели следующим образом установить “галочки”, «точки» и числа.



Затем выполнить команду “Export” и сохранить результат в файле «имя клаузы».sft. Далее на запрос программы addpros.exe “F0 файл->” просто указать имя этого файла. В результате будет получен файл «имя клаузы».san (s<peech>an<alyzer>), который следует перенести в рабочую папку с расширением “.pitch”.

Литература

Mertens, Piet (2004) The Prosogram : Semi-Automatic Transcription of Prosody based on a Tonal Perception Model // B. Bel & I. Marlien (eds.) Proceedings of Speech Prosody 2004, Nara (Japan), 23-26 March. (ISBN 2-9518233-1-2)