

Inton Trainer

SPEECH INTONATION ANALIZER&TRAINER

<https://intontrainer.by/>

ПОСЛЕДНИЕ НОВОСТИ:

(Версия: 2019-12-02)

- В режиме интонационного тренинга добавлена возможность задания желаемого количества тестовых файлов (произнесений) и получение усреднённой оценки их сходства с эталонной фразой (см. раздел 8.3)
- Добавлен вспомогательный программный модуль (Multi-Lingual Launcher) для запуска любого желаемого набора IntonTrainer-модулей (см. раздел 8.4)
- Добавлены новые IntonTrainer-модули: Belarusian IntonTrainer demo и Singing Lessons IntonTrainer demo

1

(Версия: 2019-08-22)

- Добавлено отображение положения границ голосовых участков и границ акцентных единиц фраз.

(Версия: 2019-05-04)

- Добавлена возможность анализа и сопоставления мелодических портретов фразы без необходимости предварительной ручной разметки её на пред-ядерные, ядерные и за-ядерные участки (см. раздел 7).
- Расширены возможности параметрического отображения мелодических портретов интонации (см. раздел 8.1)
- Добавлена возможность анализа и сохранения в численном виде набора просодических признаков анализируемых фраз (см. раздел 8.2).
- Увеличена точность сегментации и маркировки анализируемых сигналов.

ВНИМАНИЕ!

- ✓ ПМ «IntonTrainer» является открытой системой. Допускается модификация используемых настроек и аудио данных.
- ✓ Эталонные БД - PATTERNS -, могут быть дополнены или модифицированы пользователем в соответствии с поставленной задачей, либо формироваться заново для работы с новыми языковыми приложениями.
- ✓ Ваши вопросы, предложения и замечания просим высылать по адресу: intontrainer@gmail.com

Программный модуль «IntonTrainer» рекомендуется использовать:

- **В лингвистическом образовании**

(используется как средство визуализации интонации)

Первичное ознакомление с основными интонационными конструкциями речи, используемых в различных языках, их попарное сопоставление, изучение особенностей употребления, а также их реализации в диалоге, прозе и стихах.

- **Для самостоятельного обучения и самоконтроля**

(используется как средство интонационного тренинга)

Индивидуальный тренинг правильного воспроизведения интонации в речи и пении – изучение иностранных языков, совершенствование устно-речевых интонационных навыков в ряде профессий таких, как операторы колл-центров, дикторы радио, ТВ и др

- **В научно-практических исследованиях**

(используется как средство сравнения интонации от разных источников)

Исследование индивидуальных, эмоциональных и стилистических особенностей реализации интонации. Сравнительная оценка интонации речи и пения в норме и патологии. Оценка интонационного качества синтезированной речи.

Прототип системы «IntonTrainer» реализован в виде отдельного приложения под ОС Windows (xp, 7, 8, 10) и Linux (Debian, Ubuntu). В реализации используется библиотека Qt в качестве отладочного пользовательского интерфейса и ядра приложения, и библиотека MathGL для вывода графиков. Для реализации основного пользовательского интерфейса используется приложение, написанное на HTML/JS. Алгоритмы расчёта и обработки данных написаны на языке C с использованием библиотек GNU Scientific Library, Speech Signal Processing Toolkit (SPTK) и OpenAL.

АНАЛИЗАТОР И ТРЕНАЖЁР РЕЧЕВОЙ ИНТОНАЦИИ (руководство пользователя)

1. Назначение

Программный модуль (далее **ПМ**) «*IntonTrainer*» предназначен для анализа, экранного отображения и сопоставления интонационных контуров частоты основного тона (ЧОТ) эталонной и произносимой фраз, а также численной оценки их интонационного сходства. Оценка интонационного сходства осуществляется на базе представления интонации в виде **Мелодических Портретов (МП): универсальных – УМП** (при наличии специальной разметки анализируемых фраз), либо частично **нормированных – НМП** (при отсутствии специальной разметки). Настоящая версия ПМ обеспечивает также получение численных значений различного рода просодических признаков анализируемых фраз, описанных ниже.

2. Начало работы

Начальное окно ПМ, открывающееся после запуска программы, показано на Рис. 1.

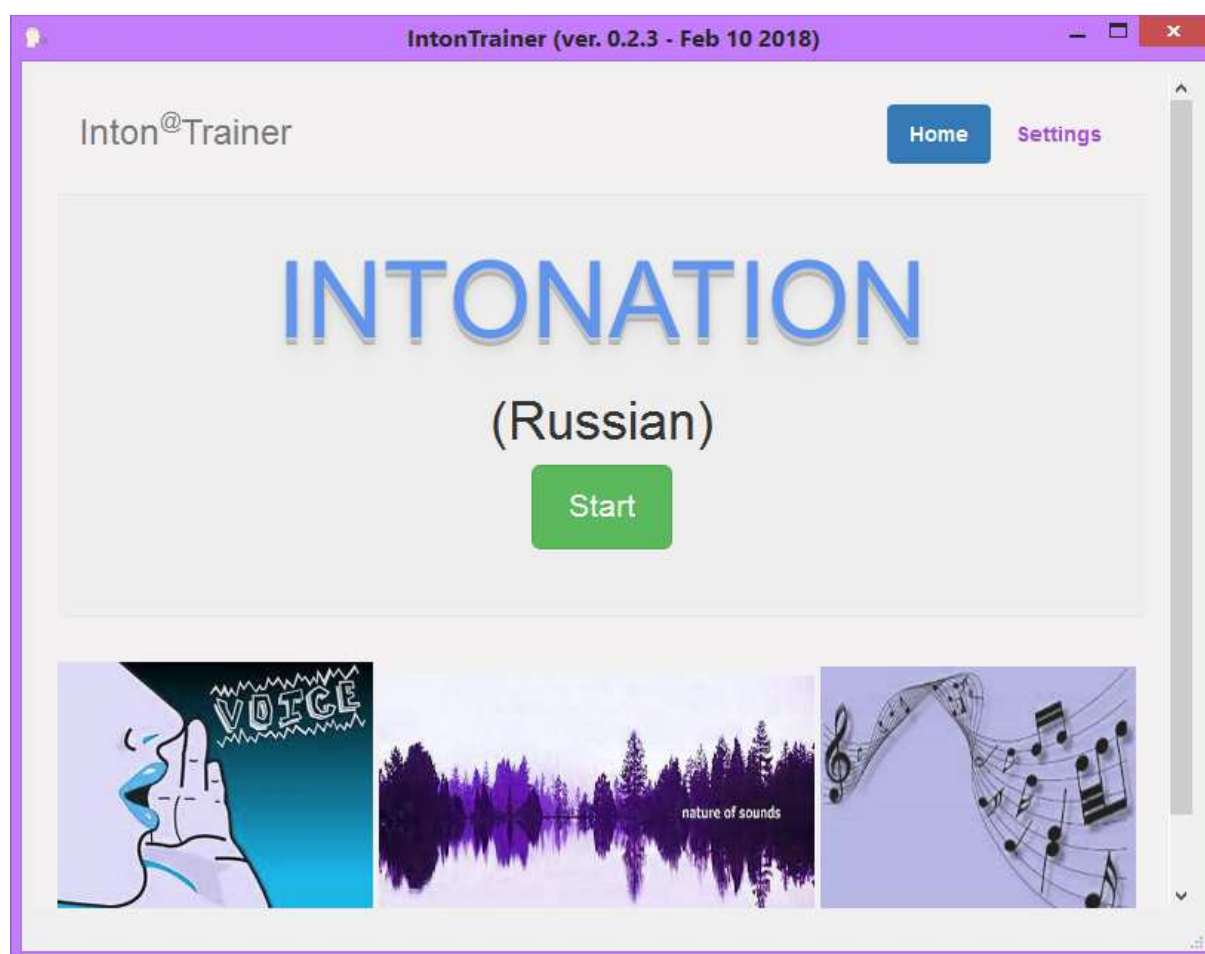
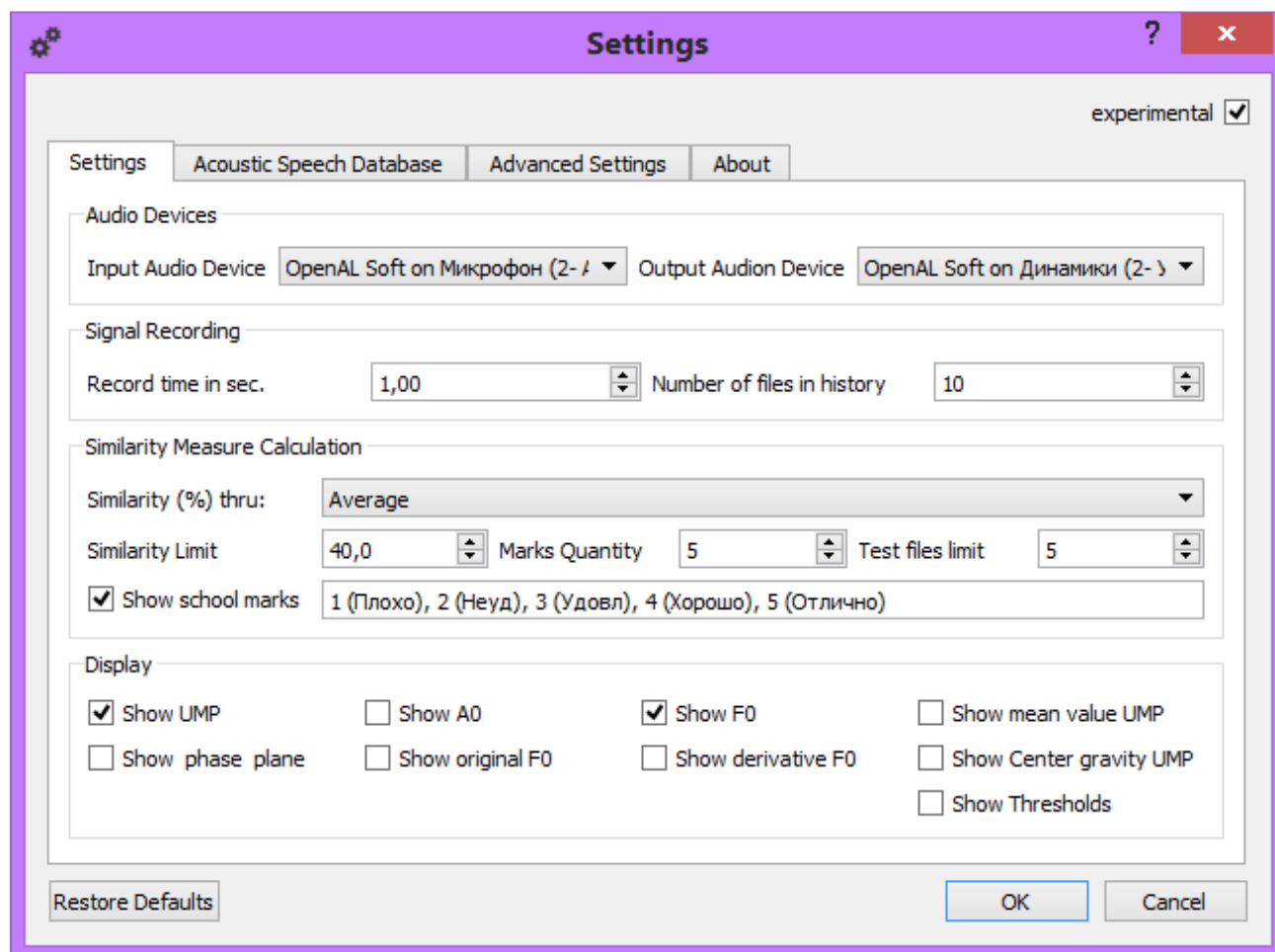


Рис. 1. Начальное окно

Перед началом работы возможен предварительный просмотр настроек **ПМ** и, при необходимости, их коррекция.

При нажатии кнопки – **Settings** – открывается начальное окно настроек ПМ (см. рис. 2), предоставляющее широкие возможности для изменения и установки желаемых персональных настроек. Возвращение к исходным настройкам осуществляется нажатием кнопки **Restore defaults** (см. левый нижний угол на рис. 2).



4

Рис. 2. Окно настроек

В этом окне настроек в разделе **Audio Devices** пользователь может выбрать тип используемых аудио устройств.

В разделе **Signal Recording** имеется возможность выбрать длительность записи сигнала с микрофона – **Record time in sec.** При этом обеспечивается запись в течении **N** секунд + длительность выбранного эталона. Кроме того, имеется возможность сохранения в папке “**RECORDS**” указываемого количества записываемых с микрофона фраз – **Number of files in history**,

В разделе **Similarity Measure Calculation** имеется возможность выбрать метод оценки близости УМП (НМП) интонации произнесённой фразы к эталонной, с использованием различных мер сходства - **Similarity (%) thru**:

- 1) взаимной корреляции – **Correlations (%)**,
- 2) среднего значение векторного расстояния – **[1 - Average Distance] (%)**,
- 3) максимума локального расстояния - **[1 - Maximum Local Distance] (%)**,

4) среднего значения 3-х вышеуказанных оценок сходства – **Average (%)**.

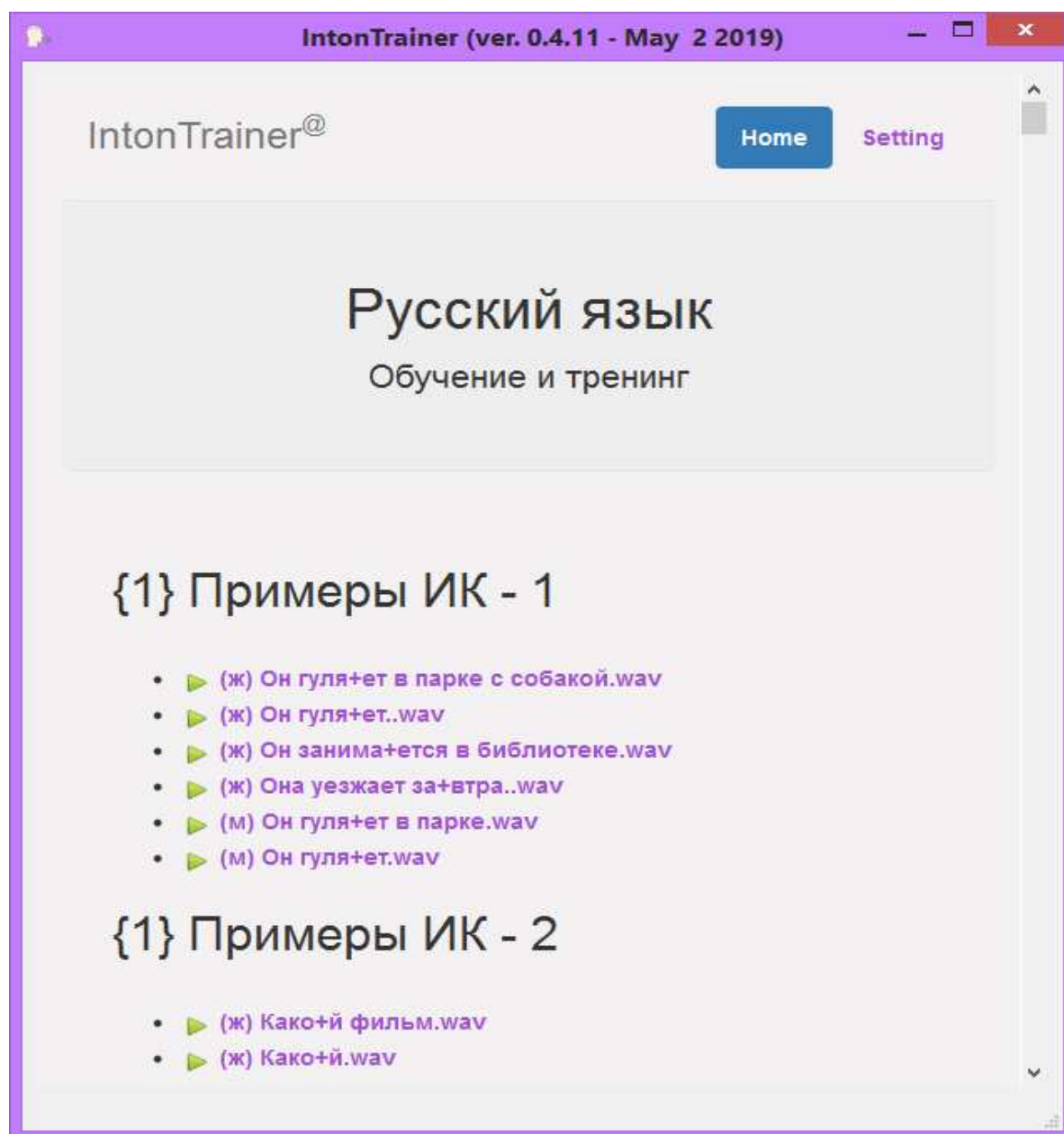
Выбранный метод вычисления интонационного сходства будет использован затем при вычислении балльной (школьной) оценки интонационного качества произнесённой фразы. Для этого делается отметка в окне **Show school marks**. Для выбранного метода экспериментальным путём определяется минимальный порог меры сходства – **Similarity Limit**, который соответствует наихудшей балльной оценке. Общее количество используемых баллов задаётся желаемым их числом – **Marks Quantity**. Путём заполнения раздела – **Mark labels** – задаются цифровые и словесные имена баллов.

В разделе **Display** пользователь может выбрать желаемый режим отображения результатов интонационного анализа исследуемых фраз. По умолчанию устанавливается основной режим настроек, отмеченный «галочками». Особенности использования других режимов будут рассмотрены ниже.

Два дополнительных раздела настроек (рис. 2) – **Acoustic Speech Database** и – **Advanced settings** - предназначены для экспериментов по подбору оптимального режима, осуществляемых разработчиками или «продвинутыми» пользователями программы. Кнопка – **About** открывает окно со сведениями о разработчиках и ряд других сведений.

3. Просмотр и прослушивание интонационных конструкций русской речи

При нажатии кнопки «**Start**» (рис. 1) открывается информационное окно ПМ (рис. 3)., содержащее структурированный перечень текстов эталонных фраз – примеров основных **интонационных конструкций (ИК)** русской речи, а также других примеров (см. рис. 4), хранящихся в папке “**PATTERNS**”.



6

Рис. 3. Информационное окно

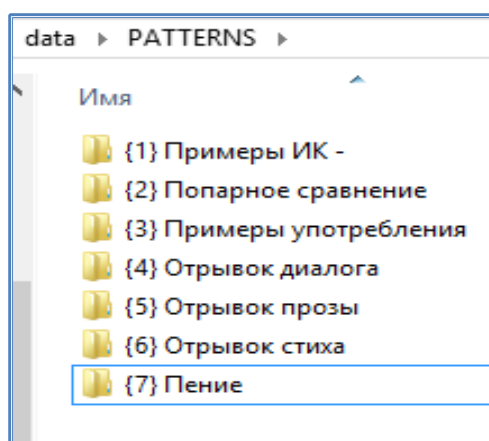


Рис. 4. Содержимое папки "PATTERNS".

Прокручивая страницу информационного окна (рис. 3), пользователь получает возможность слухового и визуального ознакомления с основными интонационными конструкциями русской речи путём выбора с помощью курсора требуемой директории, например:

{1} Примеры ИК - 5 • ► (Ж) Кому+ она только I не писа+ла!.wav

В результате такого выбора по умолчанию открывается многофункциональное **главное окно ПМ** (см. рис.5).

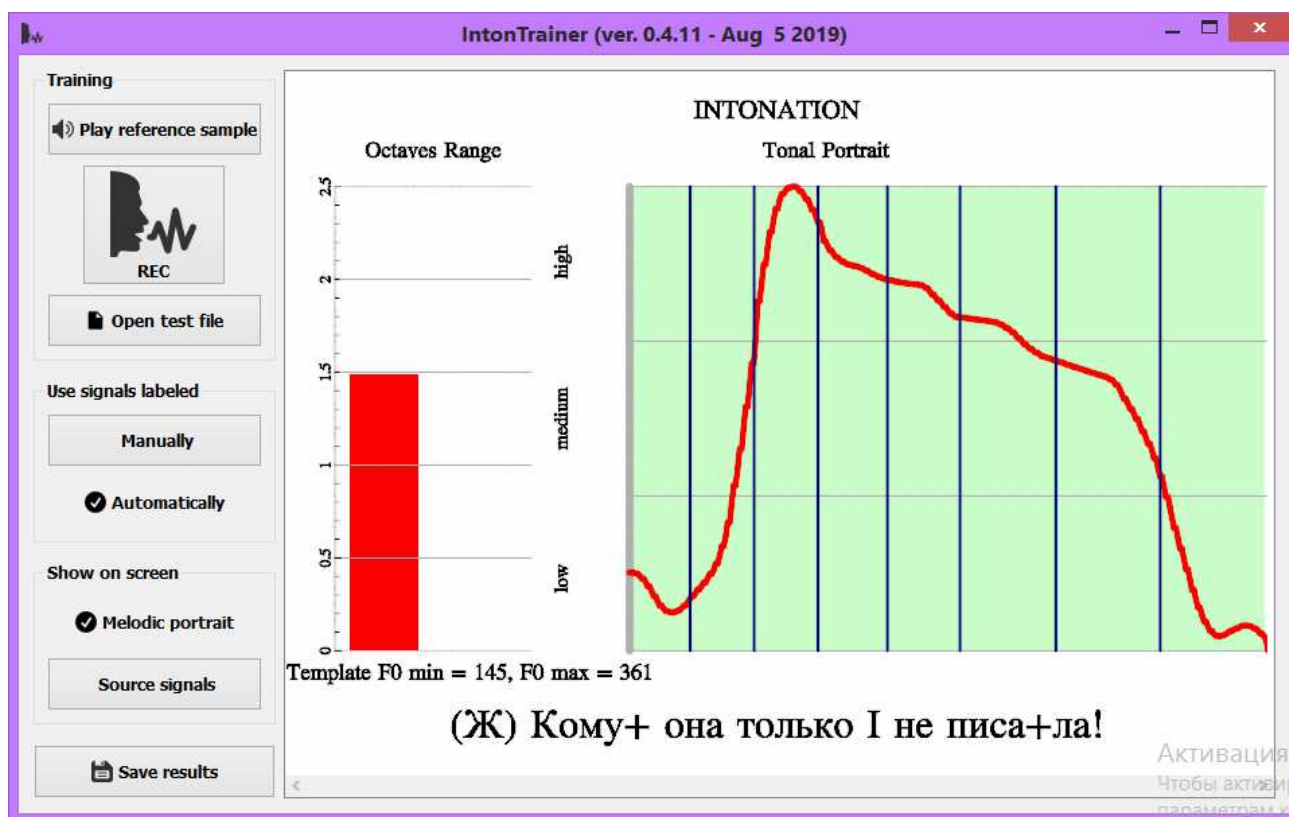


Рис. 5. Главное окно. Режим: «Automatically» – «Melodic portrait»

В **левой** части окна сосредоточены органы управления, позволяющие осуществлять прослушивание речевого сигнала, а также задавать различные режимы графического отображения результатов интонационного анализа.

Режимы графического отображения, установленные по умолчанию, отмечены «птичками», а именно:

- автоматический способ разметки сигнала: **(Use signal labeled – Automatically)**,
- вид экранного отображения результата анализа мелодического портрета: **(Show on screen – Melodic portrait)**.

В **правой** части окна в графическом виде представлены результаты анализа эталонной фразы. По умолчанию в правой части окна (рис. 5) отображается **Нормированный Мелодический Портрет (НМП)** анализируемой фразы. Под **НМП** подразумевается отображение мелодической кривой, которая нормирована относительно минимального и максимального значений ЧОТ и общей длительности голосовых участков фразы. Кривая НМП отображает динамику изменения частоты

основного тона (ЧОТ) только на озвученных (голосовых) участках фразы. Границы озвученных участков представлены вертикальными линиями. Построение кривой НМП не требует предварительной «ручной» разметки анализируемого речевого сигнала фразы. Сегментация сигнала на голосовые регионы (см. п. 7) осуществляется автоматически на основе информации о наличии периодичности в сигнале (голоса) при одновременном присутствии достаточно высокой амплитуды сигнала - **$A_0(t)$** .

Внизу под графиком кривой НМП приведены минимальное и максимальное значения F_0 фразы в Герцах, а также текст фразы, в которой ядерный гласный обозначен знаком «+».

Красный столбец (слева от мелодической кривой) отображает диапазон изменения частоты основного тона во фразе, выраженный в октавах:

$$\text{Octave's Range} = (F_{0\max} / F_{0\min}) - 1.$$

Прослушивание выбранной эталонной фразы осуществляется путём нажатия кнопки «**Play reference sample**».

На рис. 6 показан вид главного окна ПМ при выборе другого способа экранного отображения результата анализа: (**Show on screen – Source signals**).

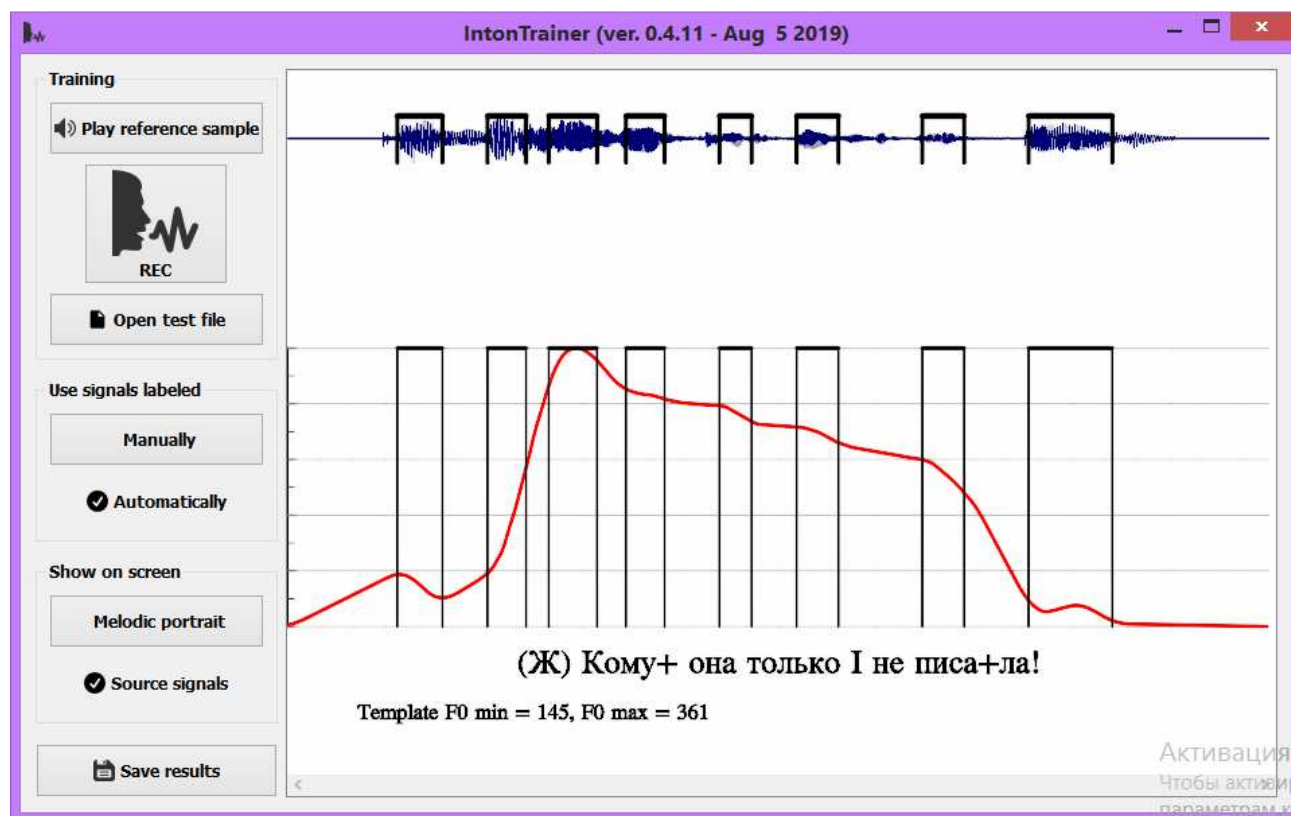


Рис. 6. Главное окно. Режим: «Automatically» – «Source signals»

В верхней части окна отображается осциллограмма сигнала фразы с указанием голосовых участков. В средней части окна изображается кривая изменения ЧОТ - $F_0(t)$, а также кривая изменения амплитуды сигнала - $A_0(t)$. Режим ПМ «**Source signals**» используется для контроля результатов анализа в тех случаях, когда возникают сомнения в правильности разметки на голосовые участки, необходимой для адекватного отображения НМП фразы.

На рис. 7 показан вид главного окна ПМ при выборе следующего способа экранного отображения результата анализа: (*Use signal labeled – Manually*) и (*Show on screen – Melodic portrait*).

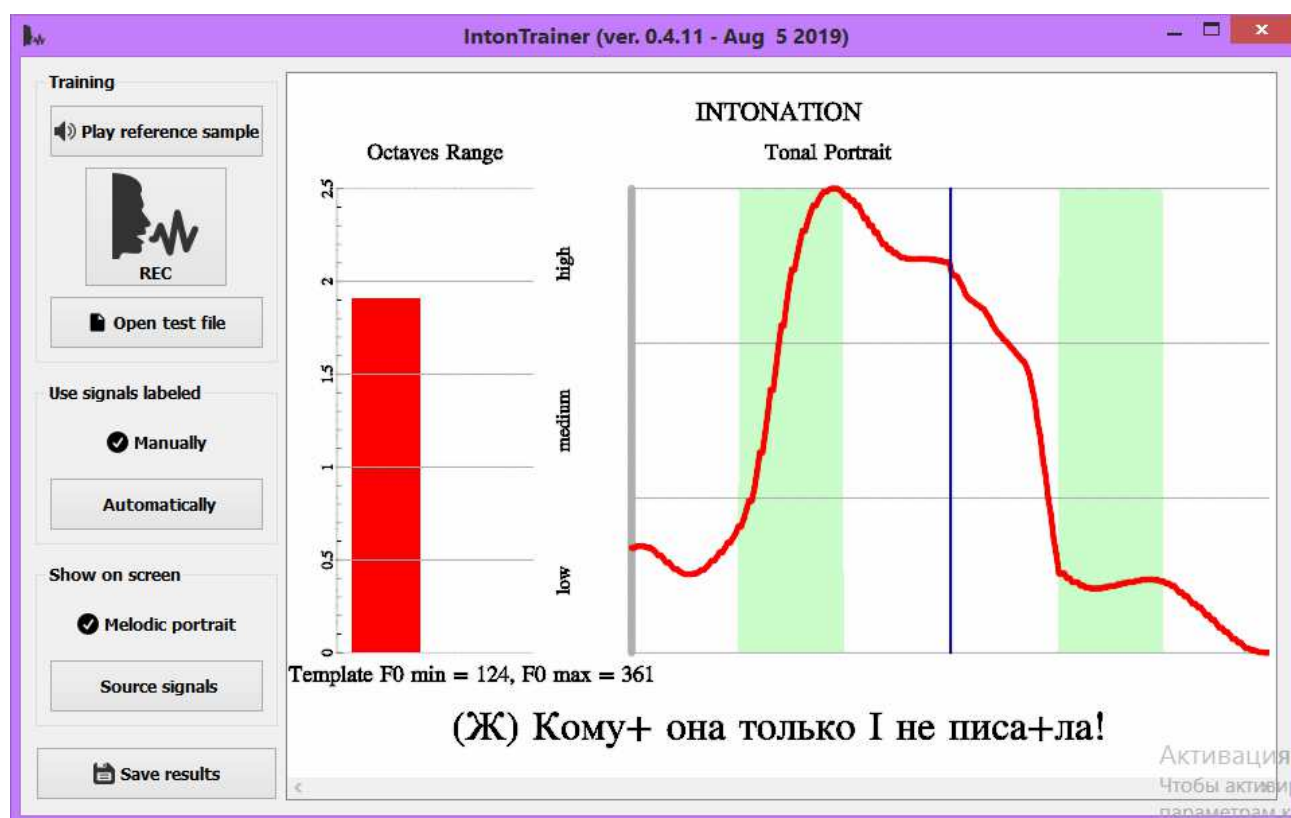


Рис. 7. Главное окно. Режим: «Manually» – «Melodic portrait»

В этом режиме кривая изменения ЧОТ в одноактной фразе отображается в виде **Универсального Мелодического Портрета (УМП)**, состоящего из трёх нормированных во времени участков: **пред-ядра, ядра и за-ядра**. Ядро помечено частыми вертикальными линиями. На рис.7 представлен результат построения УМП фразы «Куда+ же мы I теперь пое+дем?», состоящей из 2-х акцентных единиц (АЕ). На рис.7 их граница показана в виде вертикальной линии, а в тексте – значком «I».

В отличие от НМП, для построения УМП требуется предварительная ручная разметки сигнала фразы на участки **пред-ядра, ядра и за-ядра** (см. п. 7).

Внизу под графиками приведены минимальное и максимальное значения F0 для выбранной фразы, а также текст фразы, в которой ядерный гласный обозначен знаком «+».

Вид главного окна ПМ при выборе следующего способа экранного отображения результата анализа (**Show on screen – Source signals**) показан на рис. 8.

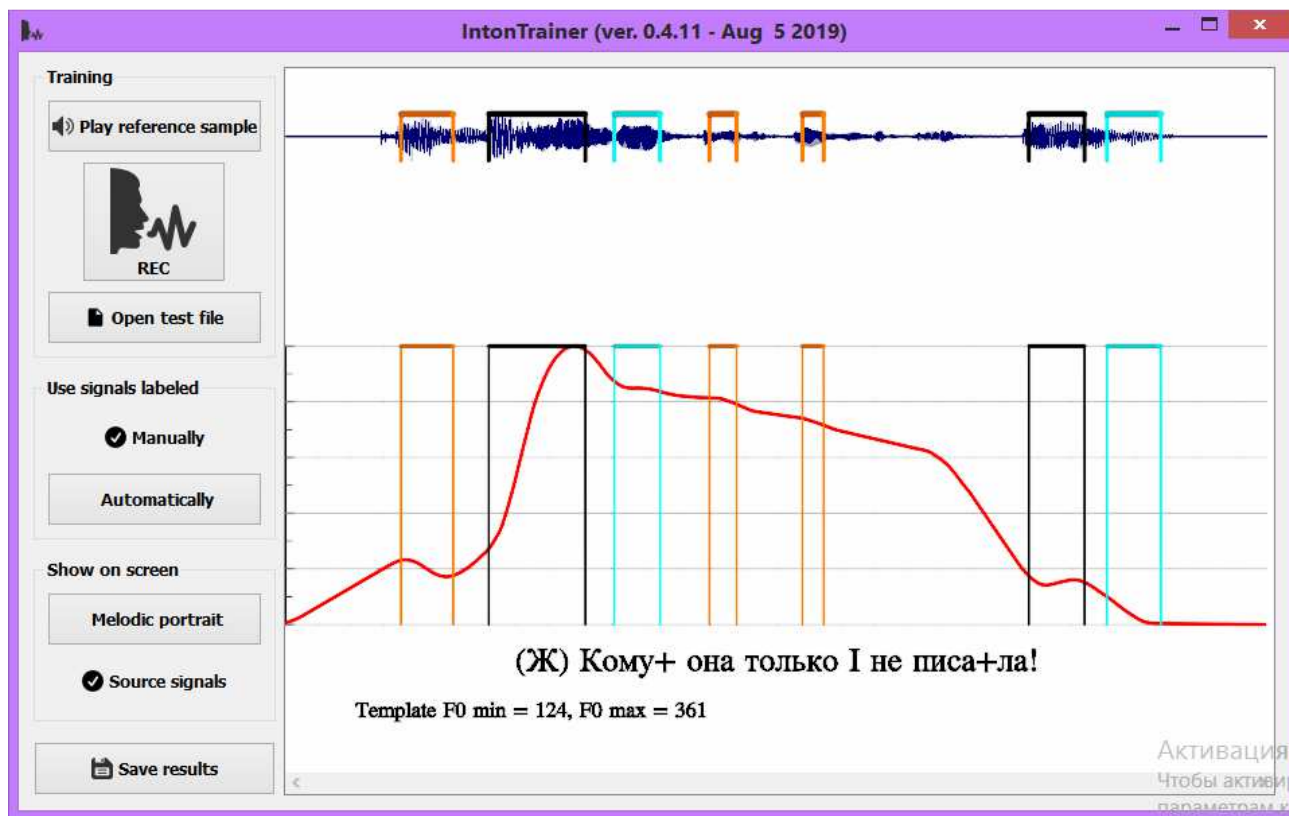


Рис. 8. Главное окно. Режим: «Manually» – «Source signals»

В верхней части окна отображается осциллограмма сигнала фразы с указанием меток пред-ядра (красная линия), ядра (чёрная линия) и за-ядра (голубая линия). В средней части окна красным цветом изображается кривая изменения F0 в реальном времени. На этой кривой указаны участки пред-ядра, ядра и за-ядра, из которых формируется УМП, показанный на рис. 7. Синим цветом изображается кривая изменения амплитуды сигнала - A0(t).

Функция **«Source signals»** используется для контроля в тех случаях, когда возникают сомнения в правильности ручной разметки для получения УМП фразы.

Листая страницу информационного окна (рис. 3) сверху вниз, пользователю предоставляется возможность слухового и визуального ознакомления с **примерами основных интонационных конструкций** русской речи (**ИК1 – ИК8**), представленных в виде УМП, а также с их попарным сопоставлением, с особенностями употребления ИК, с реализацией ИК в диалогах, прозе, стихах и при пении.

Таким образом, пользователь, просматривая последовательно описанным выше способом структурированный перечень эталонных фраз, осуществляет первичное изучение особенностей эталонной реализации **основных интонационных конструкций** русской речи.

4. Индивидуальный интонационный тренинг

Основное назначение ПМ «IntonTrainer» – предоставление пользователю возможности индивидуального интонационного тренинга при изучении русского как иностранного (РКИ), или для совершенствования уже имеющихся устно-речевых интонационных навыков русскоязычного пользователя. После того, как пользователь прослушал и просмотрел НМП или УМП выбранной фразы, он может воспользоваться внешним или встроенным микрофоном для ввода речевого сигнала своим голосом. При этом он должен нажать кнопку «**Rec**» в левой секции окна (см. рис. 7), дождаться короткого «бип-сигнала» и произнести в микрофон фразу, текст которой указан в нижней части окна на рис. 7.

После записи и обработки программой введенного речевого сигнала пользователь услышит 2-й «бип-сигнал», а изображение в окне (рис.7) заменится на изображение, показанное на рис. 9.

На рис. 9 представлены результаты интонационного анализа фразы «(ж) Он гуля+ет в парке с собакой» при выборе режима отображения: **(Use signal labeled – Manually)** и **(Show on screen – Melodic portrait)**.

В верхней части окна показаны результаты сравнения эталонной и произнесённой фраз: **Range (77%)** - близость по диапазону изменения и **Shape (84%)** – близость по форме (корреляция) сравниваемых кривых F0(t). Рядом с процентными оценками близости проставлены (при выборе режима «**Show school marks**», см. рис. 2) оценки по 5-ти бальной системе.

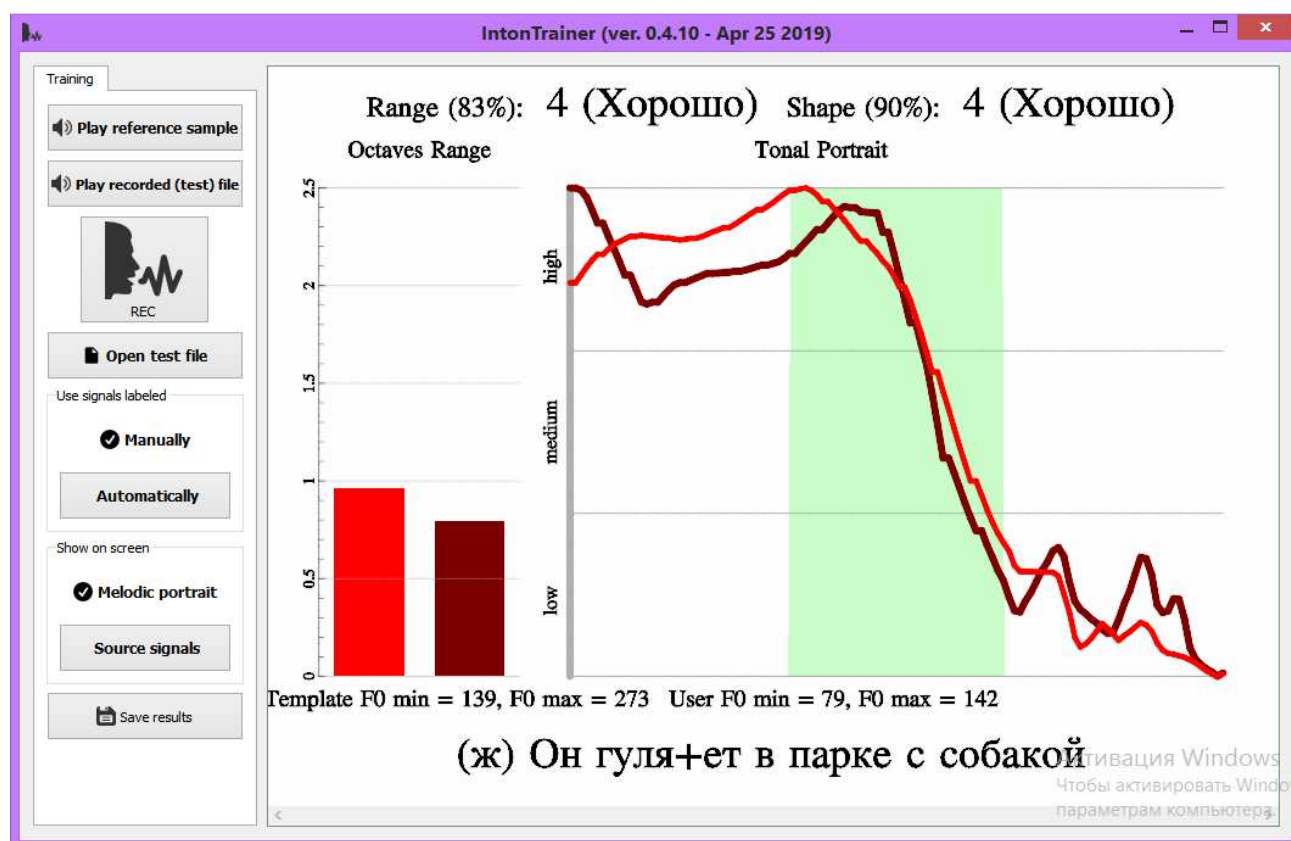


Рис. 9. Главное окно. Режим: «Manually» – «Melodic portrait»

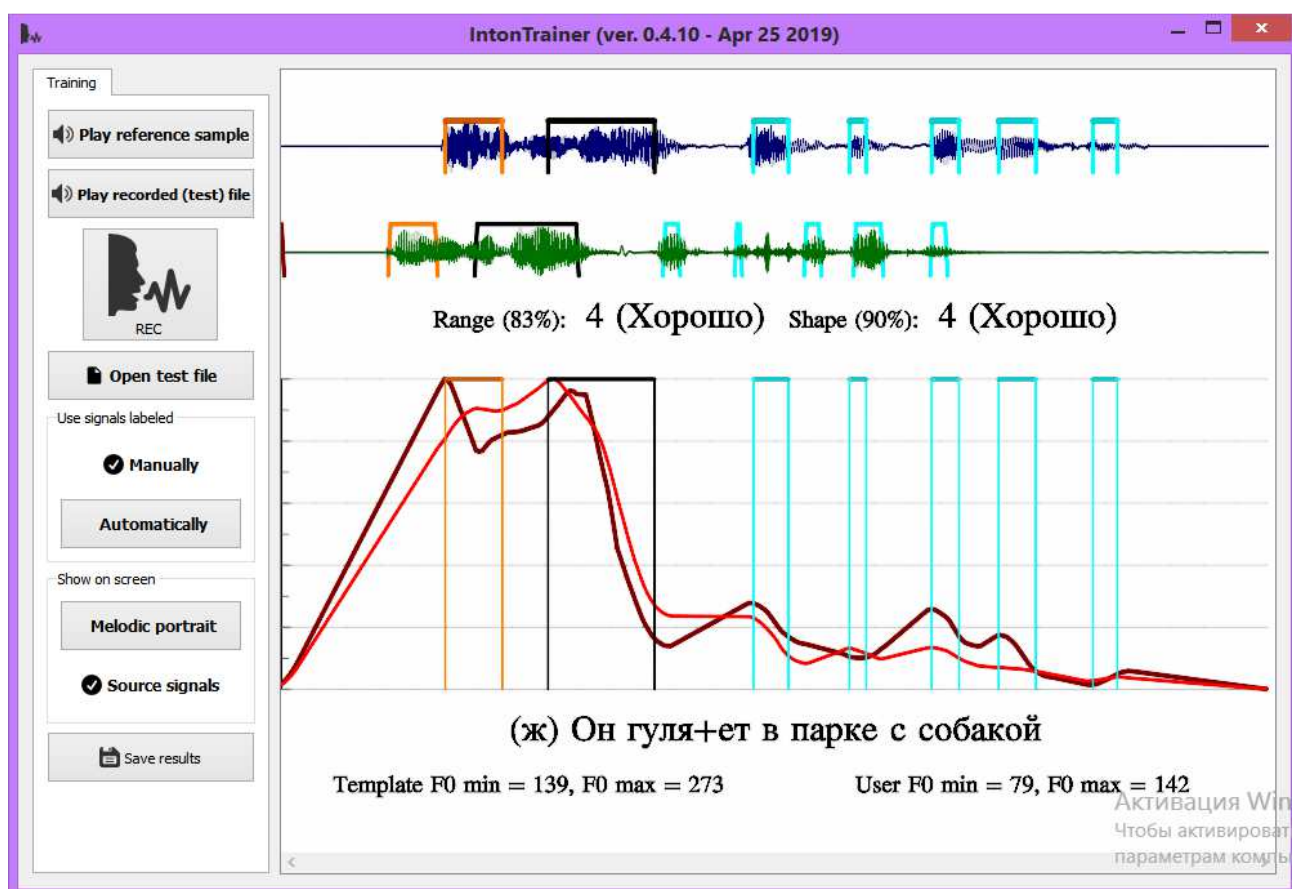
На рис. 9 **красный столбец** слева отображает диапазон изменения F0 эталонной фразы, а **коричневый** – произнесённой. Справа, **красным цветом** отображается график УМП эталонной фразы, а **коричневым** – произнесённой фразы. Внизу под графиками приведены минимальное и максимальное значения F0 эталонной и произнесённой фраз.

Прослушивание выбранной эталонной фразы осуществляется путём нажатия кнопок «**Play reference sample**», а произнесённой фразы - «**Play recorded (test) file**», находящихся в верхней части левой секции окна на рис.9.

На рис. 10 представлены результаты интонационного анализа той же фразы при выборе режима отображения: **(Use signal labeled – Manually)** и **(Show on screen – Source signals)**.

В верхней части окна отображаются осциллограммы сигналов обеих фраз (эталонной и произнесённой) с указанием меток пред-ядер (красные линии), ядер (чёрные линии) и за-ядер (голубые линии). В средней части окна изображаются совмещённые траектории изменения F0 эталонной (красная) и произнесённой (коричневая) фразы в реальном времени с указанием участков пред-ядра, ядра и за-ядра, из которых формируются УМП, показанные на рис. 9.

Информация, содержащая в этом окне, может быть полезна для контроля правильности переноса меток пред-ядра, ядра и за-ядра с эталонной фразы на произнесённую фразу. Наличие такого рода погрешностей может в значительной степени исказить действительный вид УМП фразы.



13

Рис. 10. Главное окно. Режим: «Manually» – «Source signals»

Аналогичные (но не одинаковые) результаты (см. рис. 11, 12) могут быть получены и в режиме автоматической сегментации и маркировки сигналов: (*Use signal labeled – Automatically*).

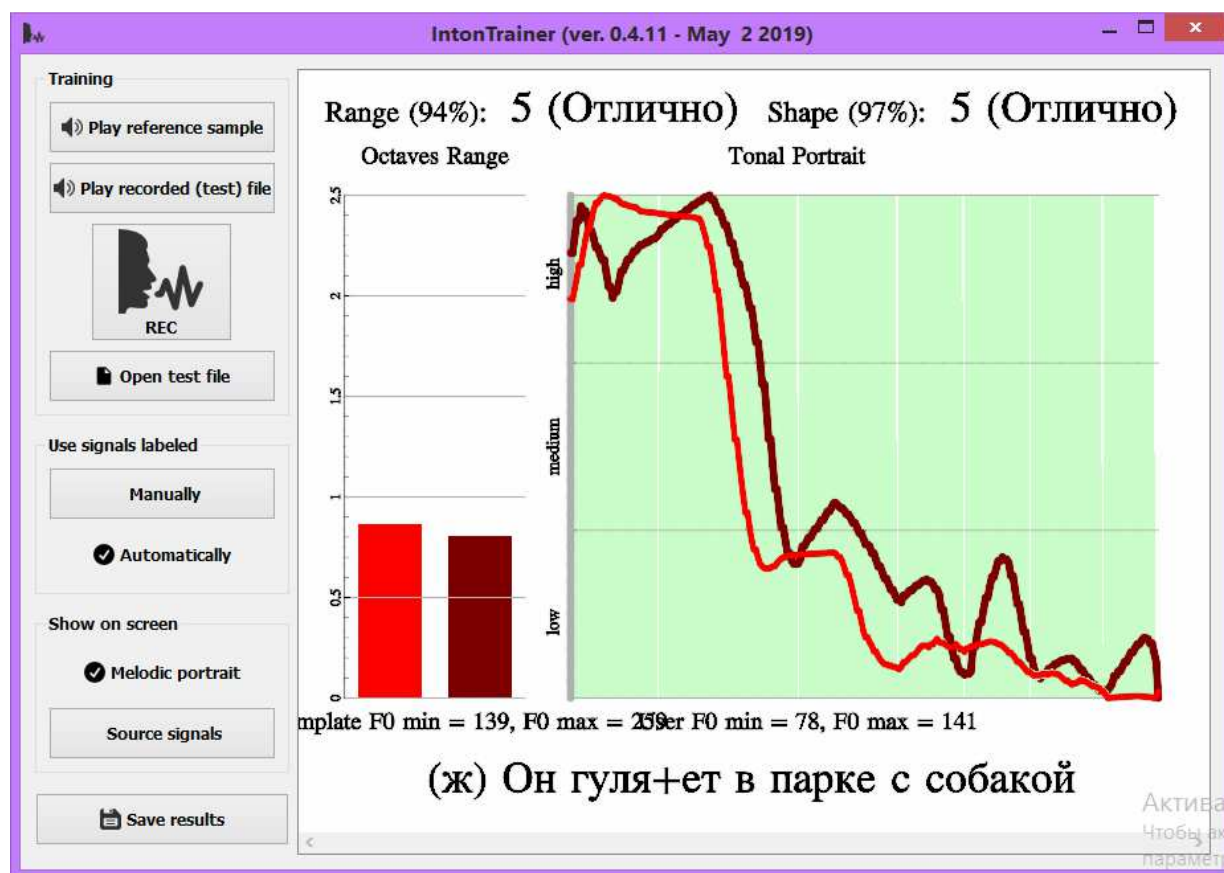


Рис. 11. Главное окно. Режим: « Automatically » – « Melodic portrait »

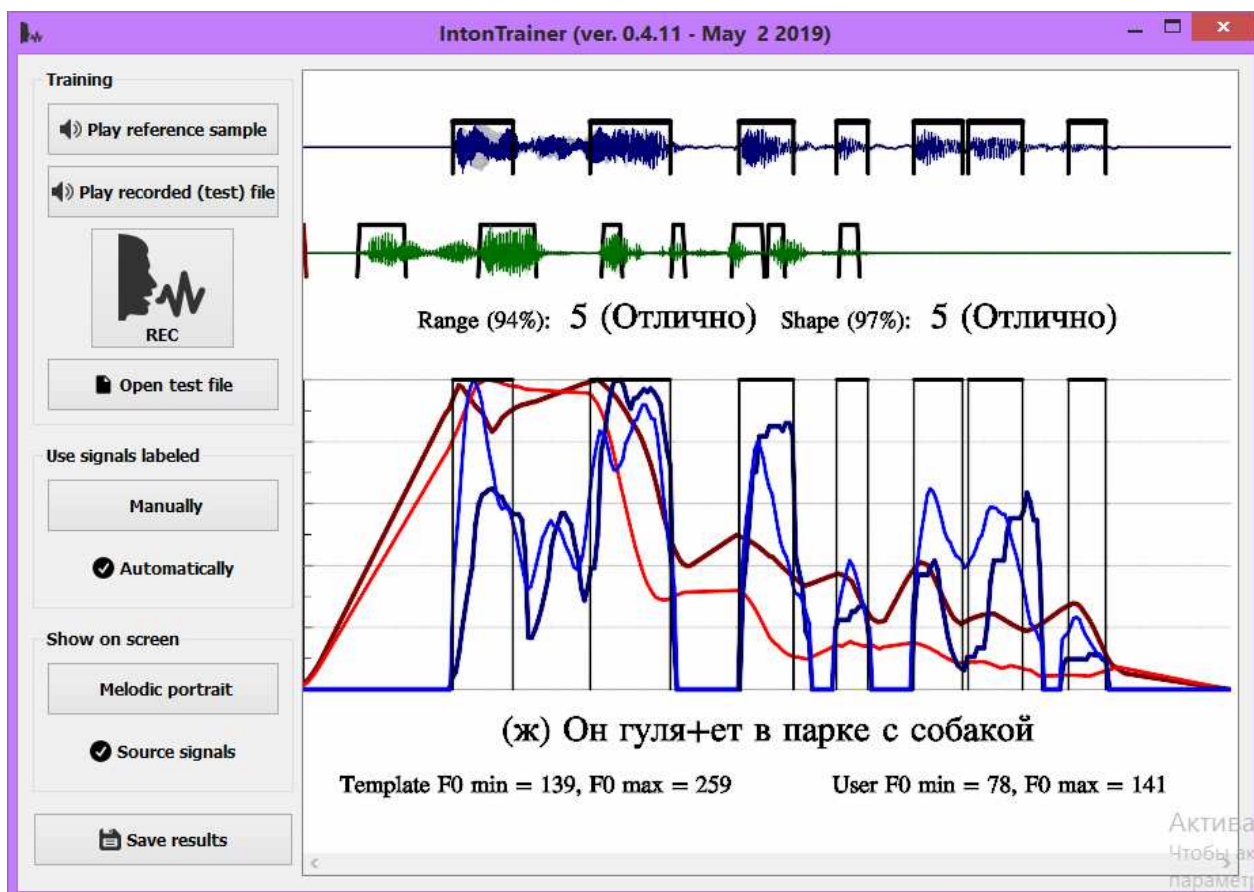


Рис. 12. Главное окно. Режим: « Automatically » – «Source signals»

Выбор автоматического либо ручного режима маркировки предоставляется пользователю. В случае выбора режима: **(Use signal labeled – Automatically)**, пользователь сможет добавлять в БД без дополнительной обработки новые эталонные данные, но при этом не гарантируется отсутствие погрешностей в маркировке голосовых сегментов. При выборе режима: **(Use signal labeled – Manually)** вероятность ошибочной сегментации и маркировки снижается и зависит только от фонетической компетенции экспериментатора (подробнее см. п. 7)..

5. Сравнение интонации фраз из различных источников

Как уже указывалось, данный ПМ может быть использован также в качестве инструмента в ряде научно-практических исследований. Например, ПМ может быть с успехом использован в экспериментально-фонетических исследованиях, в ходе которых возникает необходимость сопоставления эталонной интонации с интонацией исследуемых фраз, полученных от различного рода источников. Для этого создаётся специальная папка **“TEST”**, в которой помещаются те же фразы, что и в папке **“PATTERNS”**, но от другого источника, например, от синтезатора речи, либо фраза, произнесённая другим диктором или в другом эмоциональном состоянии. В этом случае, вместо внешнего или встроенного микрофона, используется кнопка **“Open test file”**, расположенная в левой секции Главного окна.

Пример результата сравнения интонации эталонной 2-х акцентной фразы «**Какой сегодня день!**», произнесённой диктором-женщиной (**красная линия**) и диктором-мужчиной (**коричневая линия**), показан на рис. 13.

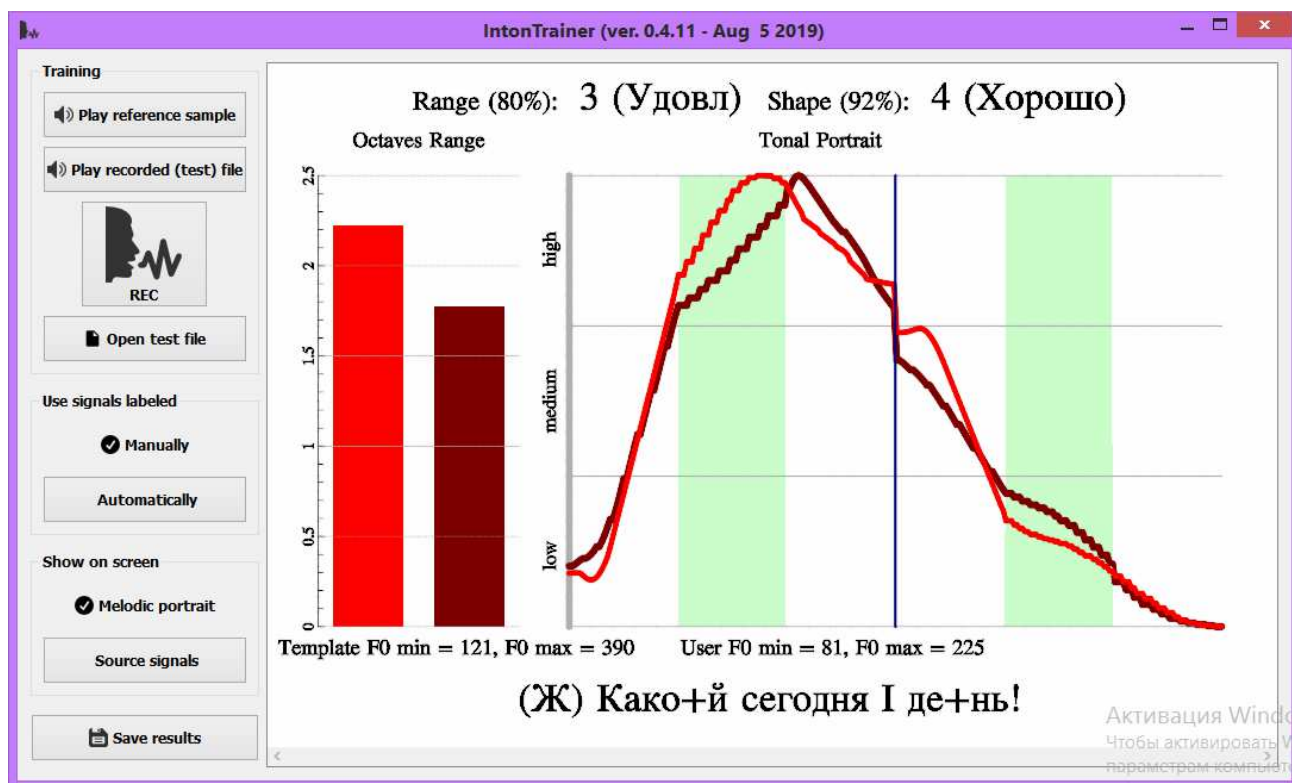


Рис. 13. Результат сравнения УМП 2-х дикторов: женщины и мужчины

В верхней части окна показаны результаты сравнения этих фраз: **Range (79%)** - сходство по диапазону изменения F0 и **Shape (93%)** – сходство по форме траектории F0.

6. Дополнительная настройка – «Acoustic Speech Database»

Программный комплекс «**IntonTrainer**» является открытой системой и допускает её всевозможные модификации. В первую очередь это относится к используемому набору эталонных данных, который может дополняться или формироваться заново в соответствии с поставленной задачей. При нажатии кнопки «**Acoustic Speech Database**» (см. рис. 2) открывается окно (рис. 14), содержащее структурированный перечень эталонных фраз.

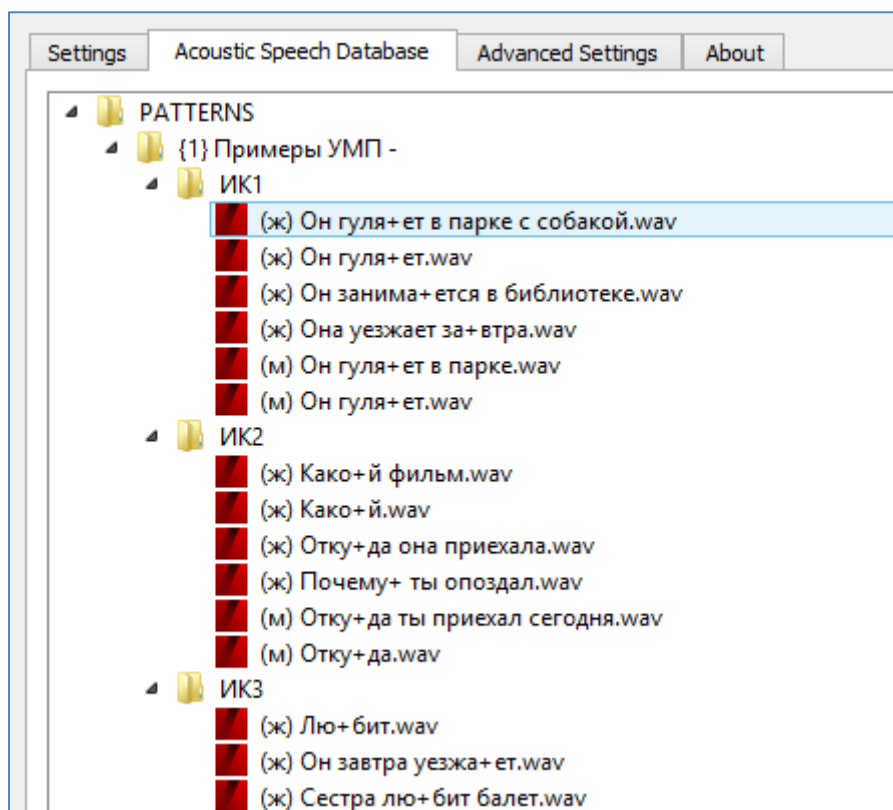


Рис. 14. Окно отображения структурированного перечня эталонных фраз

17

По двойному клику мыши на этом перечне открывается файл или окно проводника. При правом клике мыши по папке открывается выпадающее меню (рис. 15):

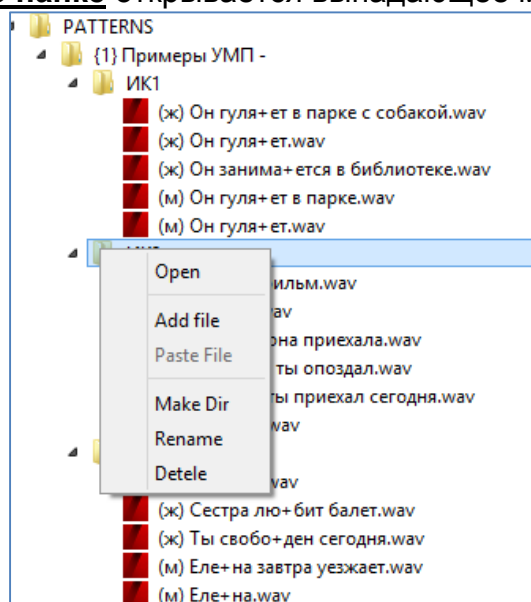


Рис. 15. Выпадающее меню при правом клике мыши по папке ИК2

Выбранную папку можно открыть (**Open**), добавить извне в эту папку файл (**Add file**), создать новую папку (**Make Dir**), переименовать (**Rename**) или стереть (**Delete**) папку.

При правом клике мыши по файлу открывается выпадающее меню (рис. 16):

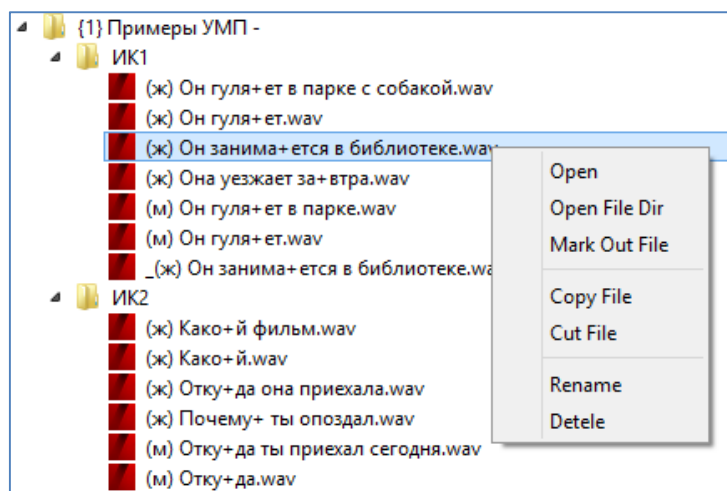


Рис. 16. Выпадающее меню при правом клике мыши по файлу:
 {(ж) Он занима+ется в библиотеке.wav}

В этом окне можно открыть любой файл (**Open**), либо открыть всю директорию (**Open File Dir**), а также **скопировать, вырезать, переименовать** или **удалить**. Таким образом, используемый набор эталонных данных (акустическая БД) может произвольно дополняться или формироваться заново в соответствии с поставленной задачей. Кроме того, используя (**Mark Out File**), предусмотрена возможность автоматической разметки любого выбранного файла на голосовые участки. В результате генерируется новый размеченный на голосовые участки файл с тем же именем, но с добавлением в начало имени знака подчёркивания { **_**(ж) Он занима+ется в библиотеке.wav }.

18

7. Просодическая разметка эталонных фраз акустической БД

Важным фактором в формировании акустической БД эталонных фраз является их просодическая разметка на **голосовые участки** (регионы) с указанием их принадлежности к пред-ядру, ядру или за-ядру. Эта операция может осуществляться «вручную» опытным фонетистом с использованием стандартной прикладной программы «**Sound Forge**». Речевой сигнал фразы записывается в «**wav**» формате с дискретизацией 8 кГц, 16 бит и размечается на регионы **Р (пред-ядро), N (ядро) Т (за-ядро)** так, как это показано на рис. 17 для одно-акцентной фразы: «**Он гуля+ет**», произнесённой женским голосом. Результат построения УМП этой фразы, полученный на основе такой разметки, представлен на рис. 18.

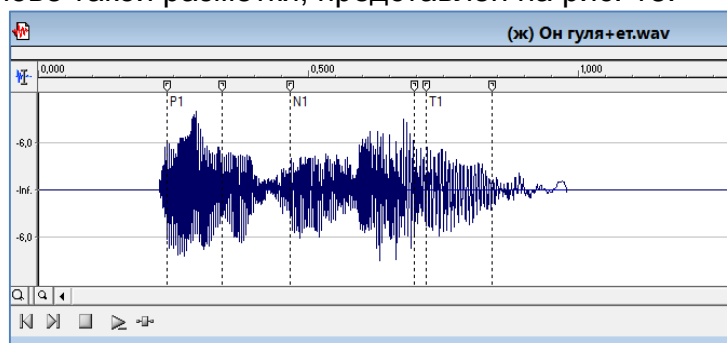


Рис. 17. Примеры разметки одно-акцентной фразы: «(ж) Он гуля+ет»

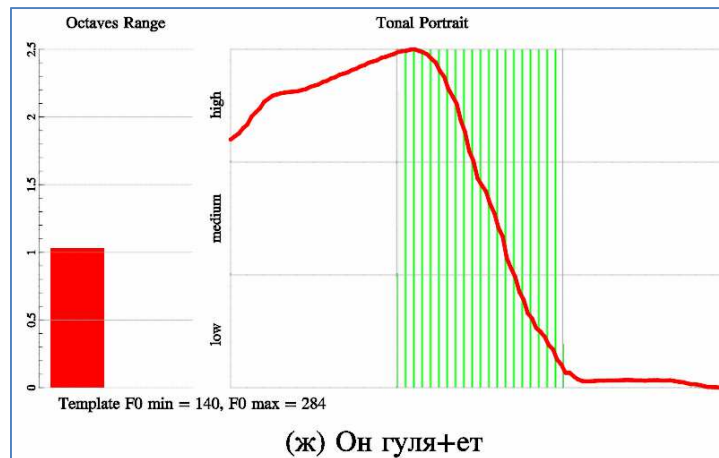


Рис. 18. Результат построения УМП фразы: «(ж) Он гуля+ет»

На рис. 19 приведен пример разметки 2-х акцентной фразы: «(Ж) Како+й сегодня де+нь!», а на рис. 20 показан результат построения УМП этой фразы, полученный на основе такой разметки.

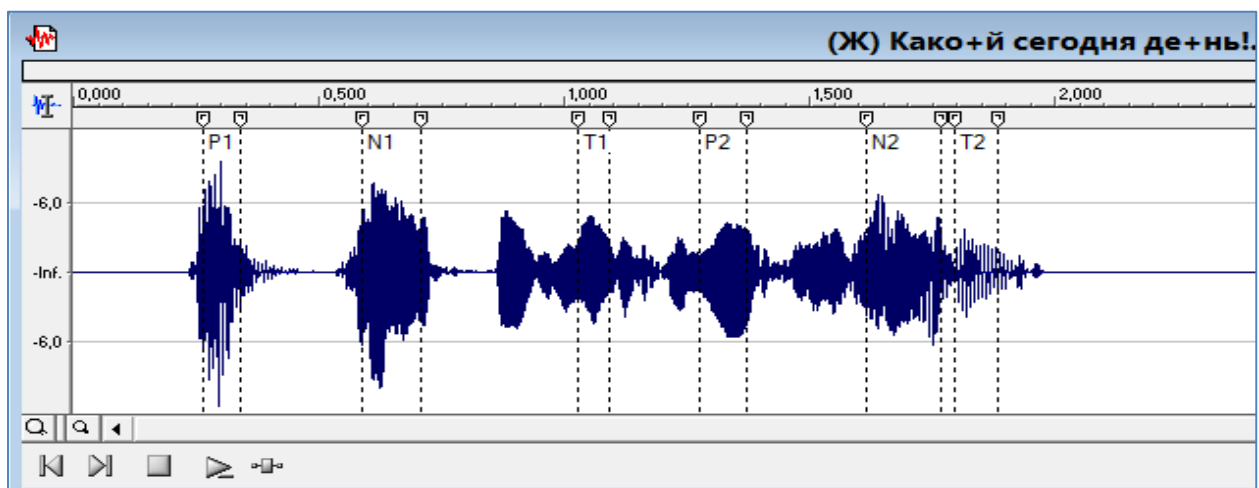


Рис. 19. Пример разметки 2-х акцентной фразы: «(Ж) Како+й сегодня де+нь!».

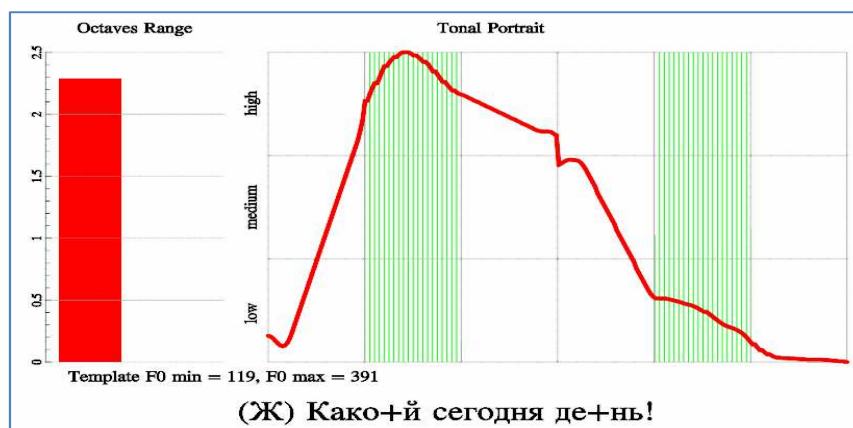


Рис. 20. Результат построения УМП фразы: «(Ж) Како+й сегодня де+нь!».

Если, по каким-либо причинам в определении положения элементов акцентной структуры фразы (P, N, T) нет необходимости или затруднительно, то всем голосовым участкам фразы можно присвоить одинаковый индекс N. В этом случае считается, что каждый из голосовых участков является «как бы» ядерным, т.е. равнозначным для описания мелодического контура всей фразы. На рис. 21 (вверху) показан пример ручной разметки осциллограммы, с использованием программы **Sound Forge**, одно-акцентной фразы «**Он гуля+ет в парке с собакой**» на (P, N, T) – регионы, а ниже на рис. 21 показан пример автоматической разметки осциллограммы, когда каждому из голосовых участков фразы присвоен одинаковый индекс N.

Создание такого файла, как уже указывалось в разделе 6, осуществляется путём выбора операции (**Mark Out File**) в дополнительных настройках – «**Acoustic Speech Database**», либо путём выбора режима: (**Use signal labeled – Automatically**). (см. раздел 3).

20

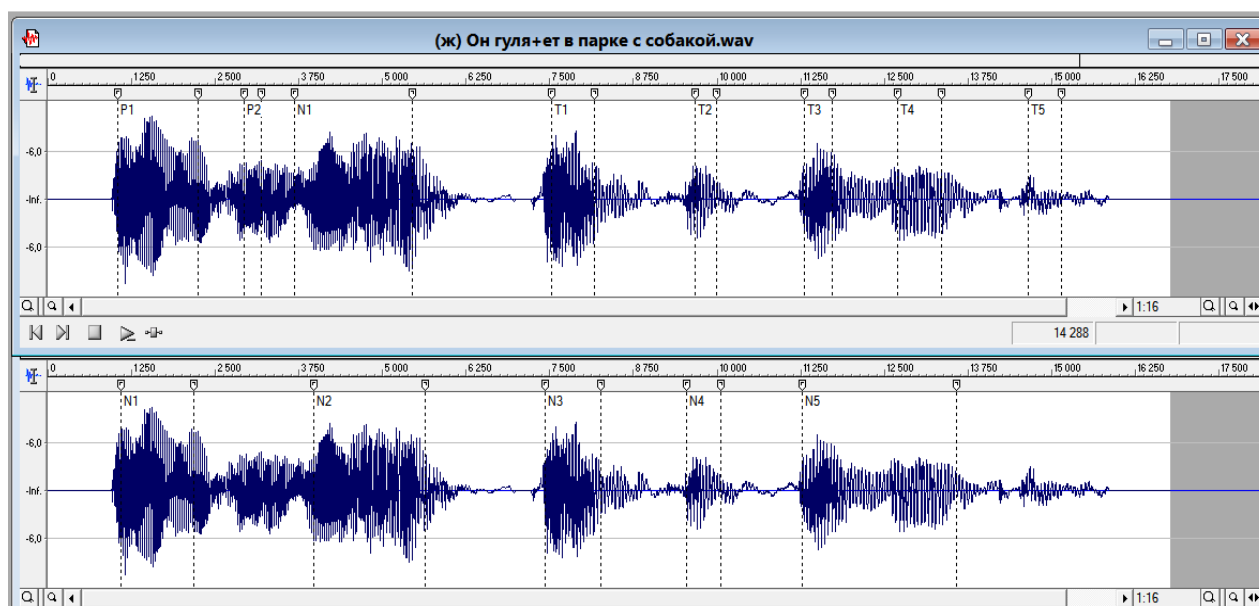


Рис. 21. Примеры ручной разметки одно-акцентной фразы (вверху) и её автоматической разметки (внизу)

Автоматическая сегментация сигнала на голосовые регионы осуществляется на основе информации о наличии периодичности в сигнале при одновременном присутствии достаточно высокой амплитуды сигнала - $A0(n)$.

Для автосегментации используется результат текущего сглаживания (интегрирования) амплитуды сигнала - $A0(n)$. Сглаживание осуществляется на интервале $(+, -) N$, где N – число отсчётов, на которых осуществляется усреднение значений кривой $A0(n)$ в каждый n -й отсчёт времени. В результате сглаживания формируются текущие значения кривой $ThRel(n)$, задающей относительный порог сегментации. Значение N задаётся в настройках как параметр $Th\ N$ (см. рис. 22). Кроме относительного порога в настройках устанавливается также абсолютный порог сегментации $ThAbs$, Причём общий порог $Th(n) = \max [Th\ abs, Th\ rel(n)]$. Графическое представление о реальной форме изменений амплитуды сигнала $A0(t)$ и о результатах сегментации сигнала фразы на голосовые регионы могут быть получены при выборе режима « *Automatically* » – « *Source signals* » и установке в разделе настроек **Display** (см. рис. 2) режима **Show A0** и **Show Thresholds** (см. рис. 22).

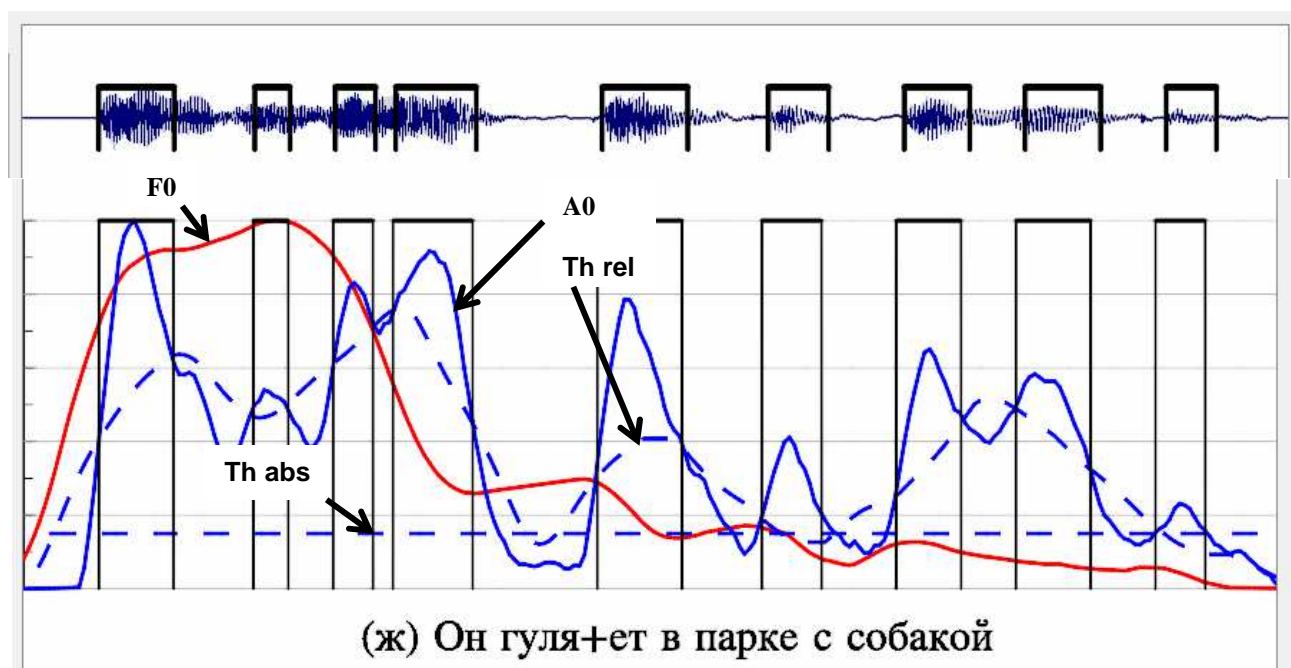


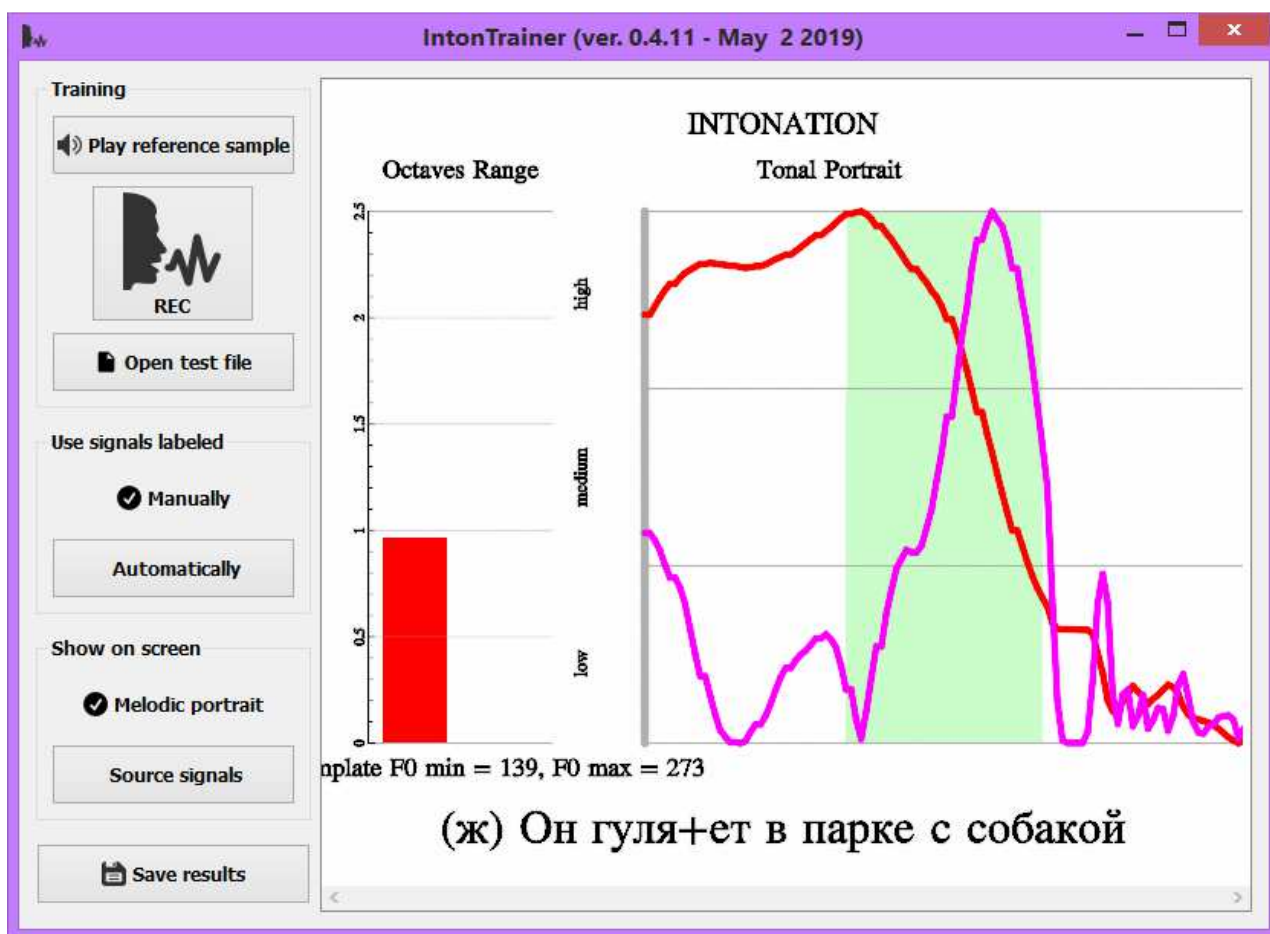
Рис. 22. Графическое представление изменений амплитуды $A0(n)$ и результатов сегментации сигнала

8. Новые дополнительные функциональные возможности ПМ «IntonTrainer»

8.1. Расширенное параметрическое отображение мелодических портретов интонации

Стандартное графическое отображение УМП, представленное ранее на рис. 7, получено при выборе в разделе настроек **Display** (см. рис. 2) **Show UMP** и **Show F0**. В случае, если будет добавлен также режим **Show derivative F0**, то получим совместное изображение УМП и его производной (см. рис. 23),

При дополнительном выборе режима **Show phase plane** получим отображение УМП и его производной на плоскости « $F_0 - dF_0$ », представленное на рис. 24.



22

Рис. 23. Совместное отображение УМП и его производной

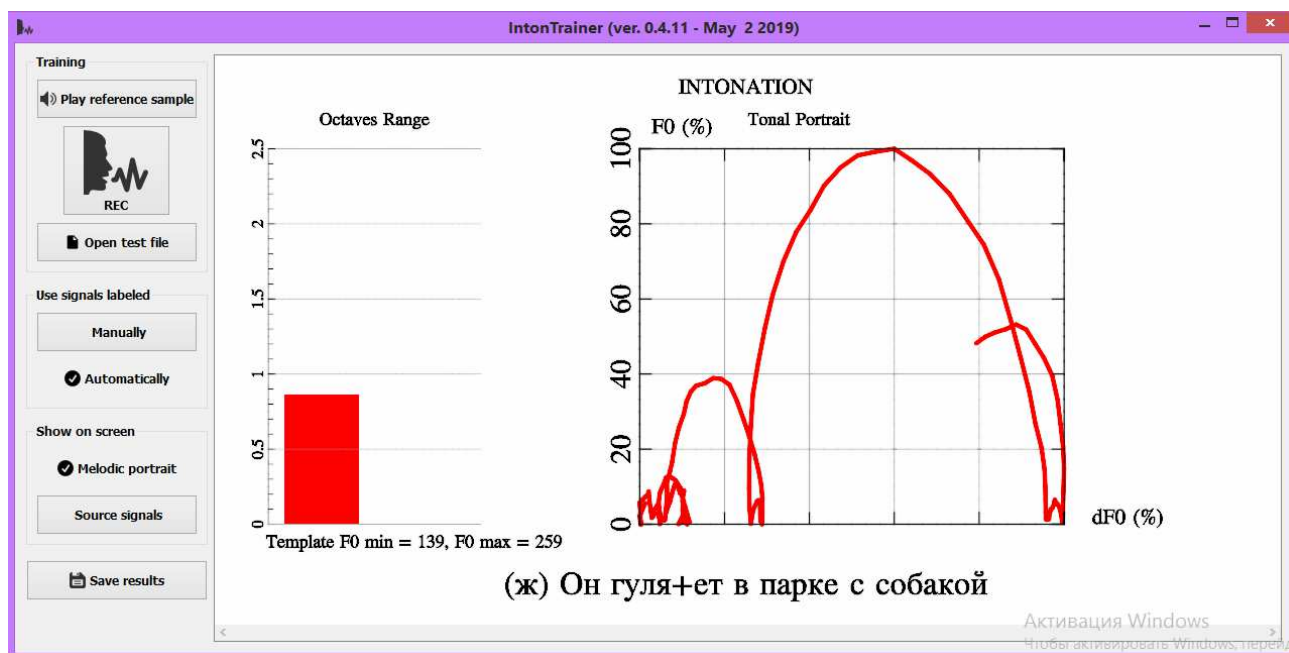


Рис. 24. Отображение УМП и его производной на плоскости «F0 – dF0»

Расширенное параметрическое отображение мелодических портретов интонации путём совместного анализа УМП и его производной позволяет учесть более тонкие различия в интонации эталонной и произносимой фраз.

Добавление возможности отображения УМП и его производной в координатах «F0 – dF0» позволяет сделать более наглядным зрительное сопоставление интонации эталонной и произносимой фраз.

23

8.2. Расчёт, визуализация и сохранение просодических признаков

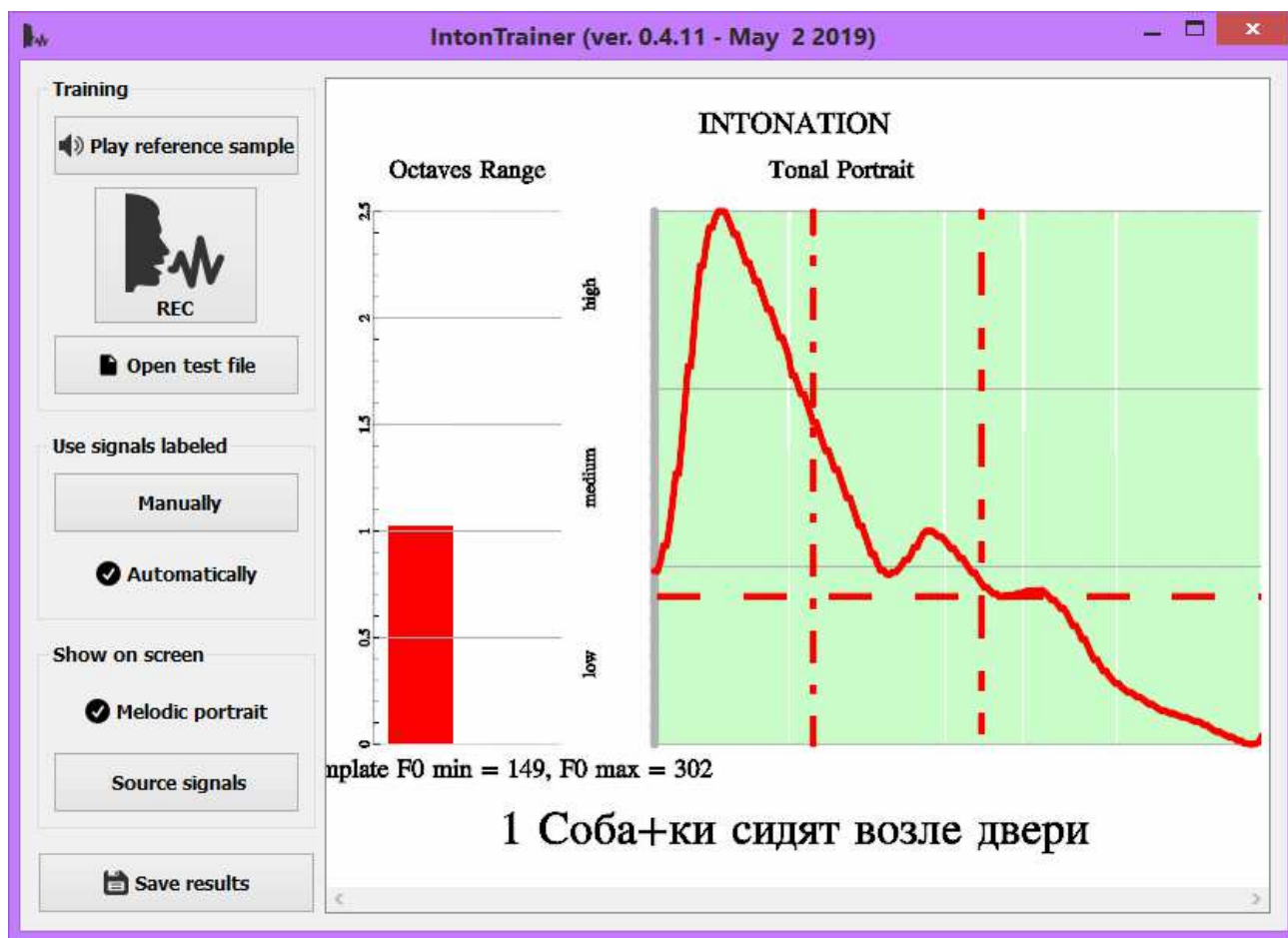
В дополнение к визуальному отображению мелодических портретов в настоящей версии ПМ добавлена возможность расчёта и сохранения набора просодических признаков анализируемых фраз. Покажем эту возможность на примере анализа признаков эмоциональной интонации при загрузке ПМ **“Emotions IntonTrainer Demo (Russian)”**.

После нажатия кнопки **«Start»** открывается **Информационное окно** Программы, содержащее структурированный перечень эталонных фраз с указанием наименования и номера БД, имени диктора, названия эмоции и текста фразы, в которой она отражена (рис. 25).



Рис. 25. Информационное окно

Путём выбора с помощью курсора требуемой директории, например:
«Minsk_Data 3 Анна-1_ 1 Нейтральность 1 Соба+ки сидят возле двери»
открывается окно, в котором отображается в графическом виде результаты
интонационного анализа этой фразы (рис. 26)




25

Рис. 26. Окно отображения кривой НМП: Анна-1 (Нейтральность)

Непрерывная кривая на рис. 26 отображает траекторию изменения ЧОТ на голосовых участках фразы и представлена в виде Нормированного Мелодического Портрета (НМП).. Горизонтальная штриховая линия показывает среднее значение НМП. Две вертикальные линии характеризуют отклонение от центра кривой НМП, т.е. её эффективную ширину на нормированной временной оси.

Высота столбика (слева от НМП) показывает диапазон изменения ЧОТ в октавах.

Внизу в левой части на рис. 26 показана кнопка управления «**Save results**», с помощью которой осуществляется сохранение данных об измеренных просодических признаках. При нажатии кнопки «**Save results**» появляется дополнительный значёк  и открывается страница в **EXCEL**:

<u>The data on the proximity</u>	<u>Proximity</u>
<u>Reference data on templates</u>	
F0 max [Hz]	302
F0 min [Hz]	149
Diapason F0	2,026845638
Register F0 [Hz]	225,5
Mean Value of the curve NMP	27,73133098
Center of the curve NMP	39,97315624
Width of the curve NMP	27,68004025
Mean Value of the Derivative curve NMP	15,78293405
Center of the Derivative curve NMP	39,2358682
Width of the Derivative curve NMP	33,25620744
Voiced Sounds Level	0,202963079
Voiced Sounds Duration	239

На странице в **EXCEL** записывается полный набор из 10-ти просодических признаков эталонной фразы (см. также табл. 1). Приведенные в таблице названия просодических признаков достаточно ясно характеризуют их физическую сущность. Полученные данные сохраняются в той же папке, где хранится исследуемая эталонная фраза.

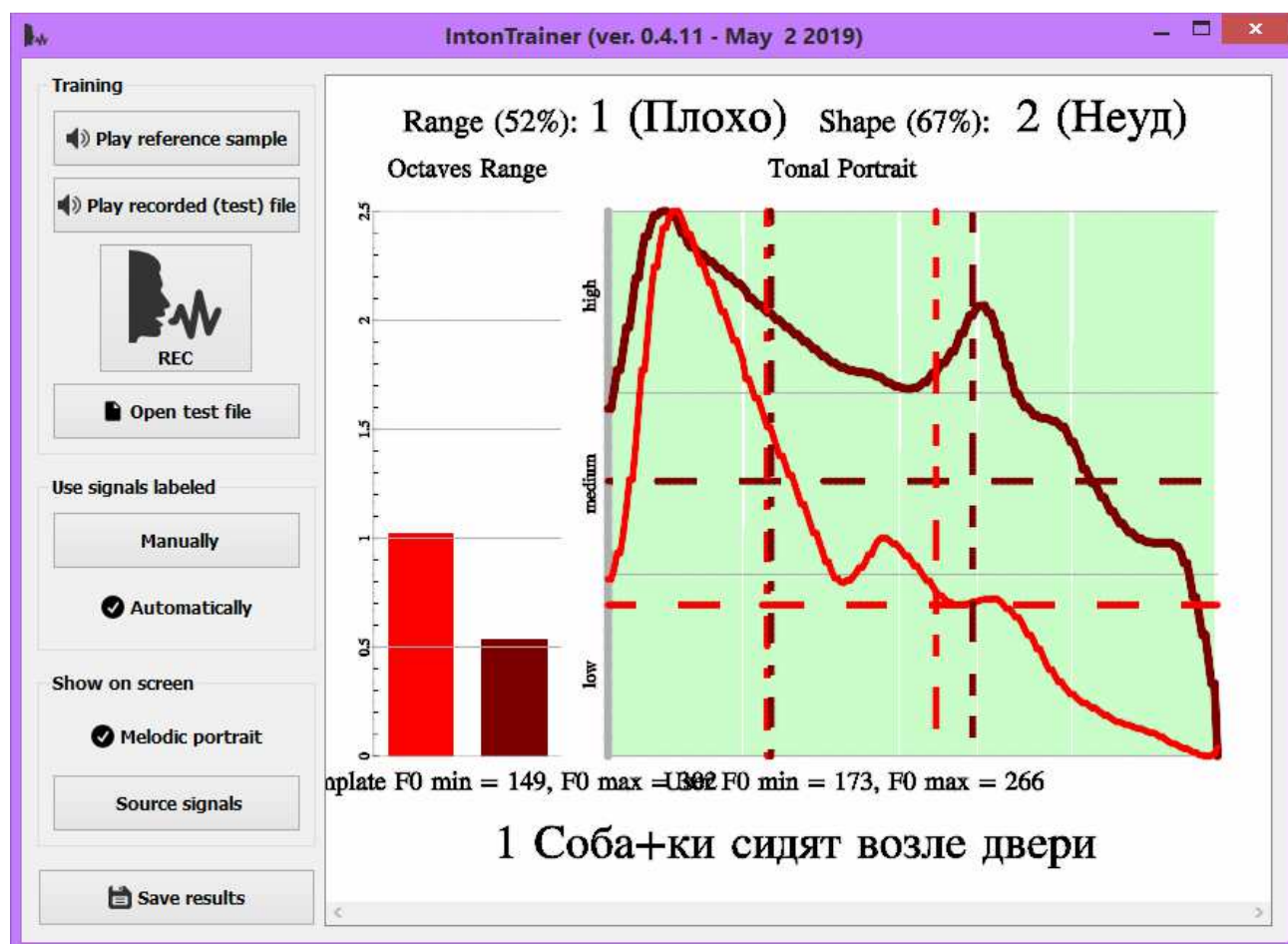
26

Таблица 1. Результаты вычисления просодических признаков: Анна-1 (Нейтральность)

№	Названия просодических признаков	Names of Prosodic Features	Results
1	Диапазон изменения ЧОТ [F0max/F0min]	Diapason F0	302,00
2	Регистр ЧОТ $[(F0max + F0min)/2]$	Register F0 [Hz]	149,00
3	Среднее значение кривой НМП	Mean Value of the curve NMP	2,03
4	Положение центра кривой НМП	Center of the curve NMP	225,50
5	Эффективная ширина кривой НМП	Width of the curve NMP	27,73
6	Среднее значение кривой d/dt (НМП)	Mean Value of the Derivative curve NMP	39,97
7	Положение центра кривой d/dt (НМП)	Center of the Derivative curve NMP	27,68
8	Ширина кривой d/dt(НМП)	Width of the Derivative curve NMP	15,78
9	Средний уровень звонких звуков	Voiced Sounds Level	39,24
10	Суммарная длительность звуков	Voiced Sounds Duration	33,26


При нажатии кнопки «**Open test file**» осуществляется вызов тестовых фраз из папки «**TEST**», в которой для сравнения просодических признаков различных эмоций помещены те же аудио файлы, что и в папке **PATTERNS**.

На рис. 27 в правой части представлен пример сравнения НМП фразы «Соба+ки сидят возле двери» эмоции «Нейтральность» (красная линия) с эмоцией «Гнев» (коричневая линия) той же фразы.



27

Рис. 27. Пример сравнения кривых НМП: Анна-1 (Гнев/Нейтральность)

После нажатия кнопки «**Save results**» и дополнительного значка  открывается страница в EXCEL, на которой записывается полный набор просодических признаков эталонной и сравниваемой фраз (см. табл. 1, 2, 3, 4).

В таблице 2 приведены результаты вычисления численных значений просодических признаков фразы, произнесённой с выражением эмоции «**Гнев**».

Таблица 2. Результаты вычисления просодических признаков: Анна (Гнев)

№	Названия просодических признаков	Names of Prosodic Features	Results
1	Диапазон изменения ЧОТ [F0max/F0min]	Diapason F0	1,45
2	Регистр ЧОТ $[(F0_{\max} + F0_{\min})/2]$	Register F0 [Hz]	228,00
3	Среднее значение кривой НМП	Mean Value of the curve NMP	54,47
4	Положение центра кривой НМП	Center of the curve NMP	46,00
5	Эффективная ширина кривой НМП	Width of the curve NMP	37,40
6	Среднее значение кривой d/dt (НМП)	Mean Value of the Derivative curve NMP	49,28
7	Положение центра кривой d/dt (НМП)	Center of the Derivative curve NMP	44,31
8	Ширина кривой d/dt (НМП)	Width of the Derivative curve NMP	50,73
9	Средний уровень звонких звуков	Voiced Sounds Level	0,19
10	Суммарная длительность звуков	Voiced Sounds Duration	226,00

В таблице 3 приведены результаты вычисления (по данным, приведенным в таблицах 1 и 2) относительных значений для просодических признаков пары фраз с эмоциями «Гнев/Нейтральность», выраженных в децибелах. Использование относительных величин позволяет осуществлять сравнение пары фраз с различными эмоциями, используя просодические признаки, выраженные в различных единицах измерения.

В таблице 4 приведен также пример результатов вычисления численных мер сходства и расстояний между 2-мя предъявленными реализациями эмоциональных фраз (в данном случае пара: (Нейтральность – Гнев).

**Таблица 3. Численные значения относительных признаков: Анна
(Гнев/Нейтральность)**

№	Названия просодических признаков	Names of Prosodic Features	Relations [dB]
1	Диапазон изменения ЧОТ [F0max/Fomin]	Diapason F0	-0,67
2	Регистр ЧОТ [(F0max +Fomin)/2]	Register F0 [Hz]	0,92
3	Среднее значение кривой НМП	Mean Value of the curve NMP	-1,59
4	Положение центра кривой НМП	Center of the curve NMP	-0,09
5	Эффективная ширина кривой НМП	Width of the curve NMP	2,38
6	Среднее значение производной кривой НМП	Mean Value of the Derivative curve NMP	0,65
7	Положение центра производной кривой НМП	Center of the Derivative curve NMP	1,41
8	Эффективная ширина производной кривой НМП	Width of the Derivative curve NMP	1,55
9	Средний уровень звонких звуков	Mean Value of Voiced Sounds Level	-0,07
10	Общая длительность звуков фразы	Total duration of phrase sounds	-0,26

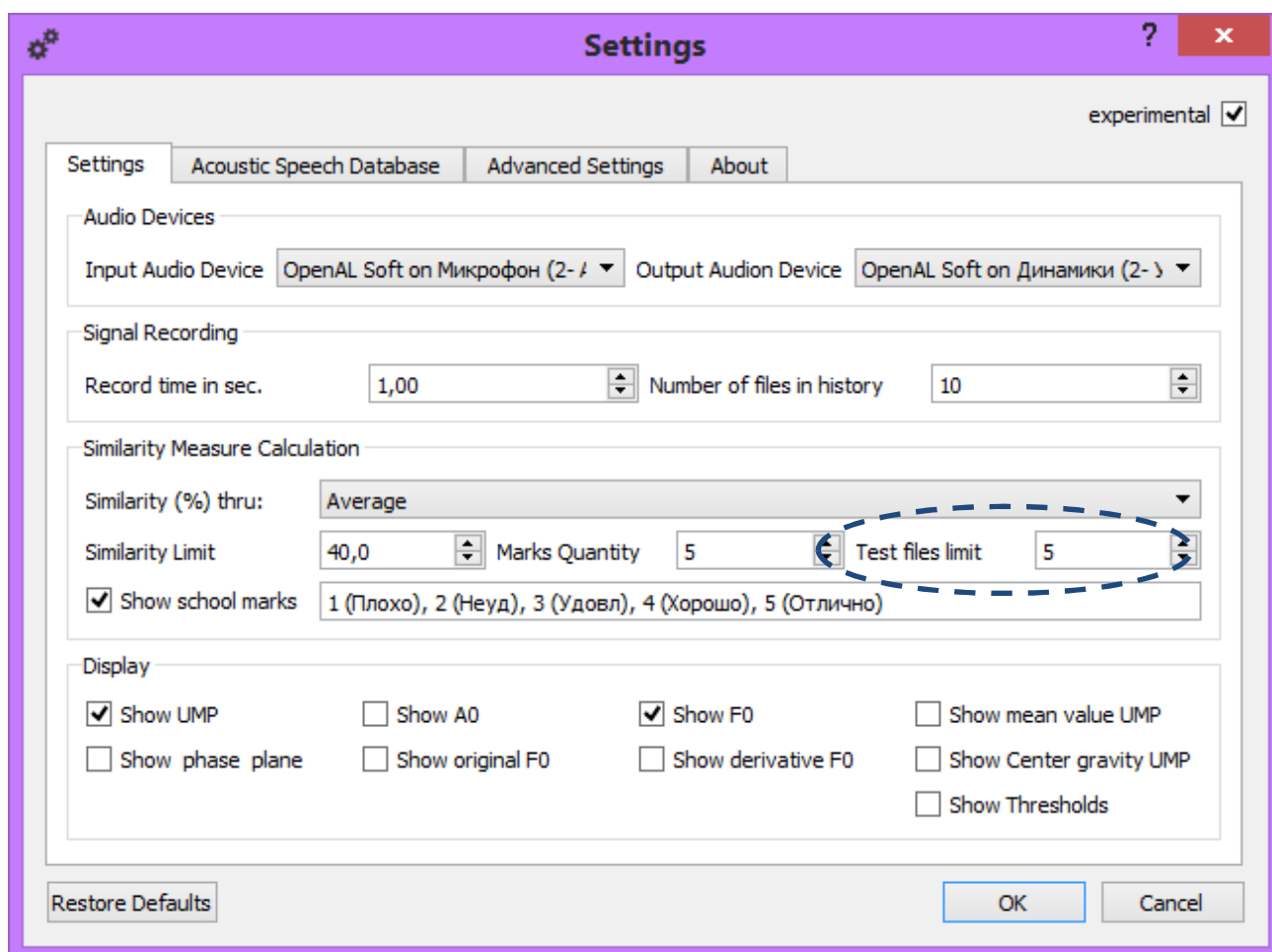
29

**Таблица 4. Результаты вычисления мер сходства и расстояний: Анна
(Гнев/Нейтральность)**

№	Способ сравнения сходства	Type of the proximity	Proximity	Distance
1	Коэффициент взаимной корреляции	Proximity thru curve "Correlation"	91	9
2	Интегральное сравнение кривых НМП	Proximity thru curve "Integral"	68	32
3	Локальное сравнение кривых НМП	Proximity thru curve "Local"	44	56
4	Среднее значение 3-х способов сравнения	Average of the three above proximities	68	32
5	Сравнение диапазонов изменения ЧОТ	Proximity thru "Range"	42	58

8.3. Усреднённая оценка сходства произносимых фраз с эталонной

В режиме интонационного тренинга добавлена возможность задания желаемого количества тестовых файлов (произнесений) и получение усреднённой оценки их сходства с эталонной фразой. Для этого в начальном окне настроек – **Settings** – в разделе - **Test files limit** – выбирается желаемое максимальное количество тестовых файлов, для которых будет вычисляться усреднённая оценка их сходства с эталонной фразой (см. рис. 28).



30

Рис. 28. Окно настроек

На рис. 29 показаны результаты вычисления усреднённых оценок сходства 5-ти произнесений фразы «Он гуляет» с эталонной.

В левой части рисунка (**Octaves Range**) графически показаны результаты вычисления сходства по диапазону изменения ЧОТ для каждого произнесения (тонкие столбцы, причём, черному цвету соответствует последнее произнесение) и результирующей усреднённой оценки (толстый столбец). Вверху показывается численная величина усреднённой оценки – **Range = 90%..**

В правой части (**Tonal Portrait**) показаны результаты вычисления УМП (или НМП) произносимых фраз (тонкие линии) и результирующая усреднённая кривая (толстая линия). Вверху показывается численная величина усреднённой оценки сходства с эталонной кривой - **Shape (85%)**.

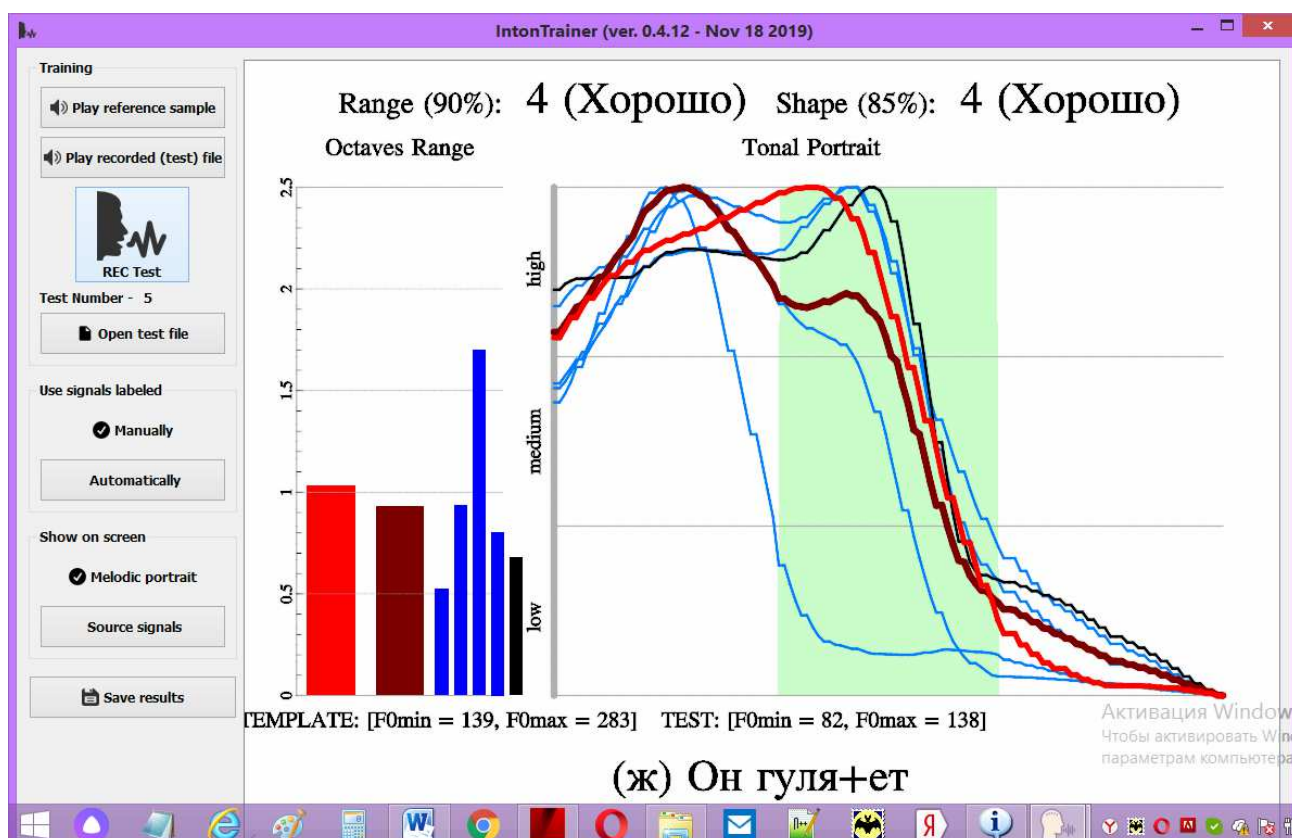


Рис. 27. Результаты вычисления усреднённых оценок сходства 5-ти произнесений

31

8.4. Сервисный программный модуль – Multi-Lingual Launcher

В настоящий момент на сайт intontrainer.by загружено 10 версий ПМ, различающихся по языковому или функциональному признаку. Сервисный программный модуль – **Multi-Lingual Launcher** – создан с целью обеспечения удобства работы пользователя с нужным ему набором ПМ **IntonTrainer**. Для реализации этой возможности пользователю достаточно скопировать нужный ему набор ПМ в папку Applications.

На рис 28 приведен вид рабочего окна при запуске программы **Multi-Lingual Launcher** в том случае, когда пользователю понадобился бы полный набор ПМ. Пользуясь этой сервисной программой, пользователь имеет возможность оперативного вызова одной или нескольких интересующих его ПМ.



32

Рис. 28. Вид рабочего окна при запуске программы **Multi-Lingual Launcher**

9. ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Узнать больше о теоретических основах разработки данного ПМ: см. дополнительную информацию, размещенную на этом сайте.

Мы также рекомендуем пользователям время от времени заглядывать на наш сайт, в котором постоянно могут появляться полезные обновления программы «IntonTrainer»