

= Inton[@]Trainer =

АНАЛИЗАТОР И ТРЕНАЖЁР РЕЧЕВОЙ ИНТОНАЦИИ (руководство пользователя)

1. Назначение

Программный комплекс «IntonTrainer» (далее «Приложение») предназначен для анализа и экранного отображения интонации (основного тона – F0) эталонных и произносимых фраз, а также их сопоставления для оценки интонационного сходства. Оценка интонационного сходства осуществляется на базе представления интонации в виде *универсальных мелодических портретов (УМП)*.

Программный комплекс рекомендуется к использованию в следующих достаточно широко востребованных приложениях:

- В лингвистическом образовании
(используется как средство визуализации интонации)

Первичное ознакомление и изучение основных интонационных конструкций (ИК) русской речи, их попарного сопоставления, особенностей употребления, а также их реализации в диалоге, прозе и стихах.

- Для самостоятельного обучения и самоконтроля
(используется как средство интонационного тренинга)

Индивидуальный тренинг произношению ИК русской речи: изучение русского как иностранного (РКИ), 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 в норме и патологии. Оценка интонационного качества синтезированной речи.

2. Графический интерфейс пользователя

2.1 Начальное окно Приложения, открывающееся после запуска программы, представлено на рис. 1.

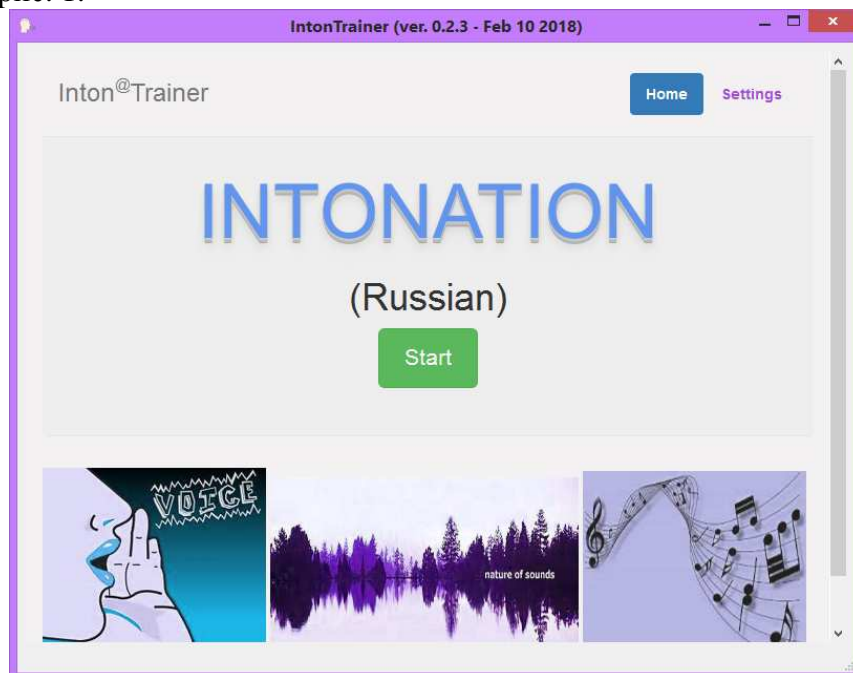


Рис. 1. Начальное окно

После нажатия кнопки «**Start**» открывается **главное окно** Приложения, содержащее структурированный перечень эталонных фраз (рис. 2)

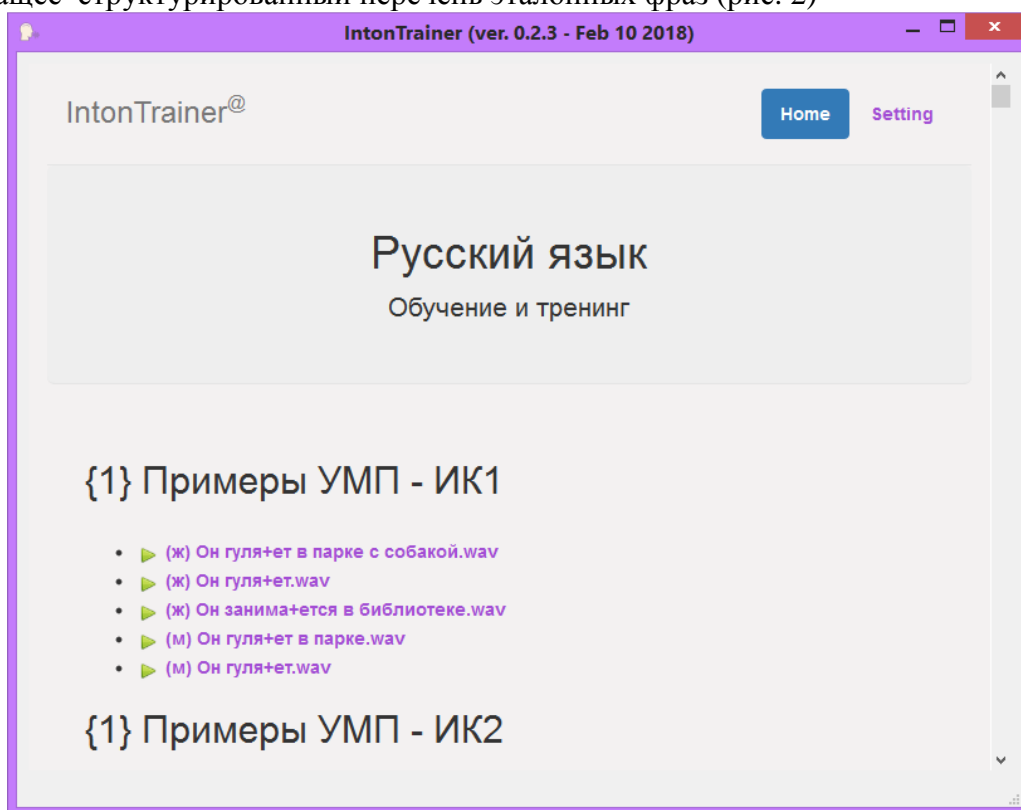


Рис. 2. Главное окно

Перед началом работы возможен предварительный просмотр настроек Приложения и их коррекция. При нажатии кнопки – **Settings** – открывается начальное окно настроек (см. рис. 3).

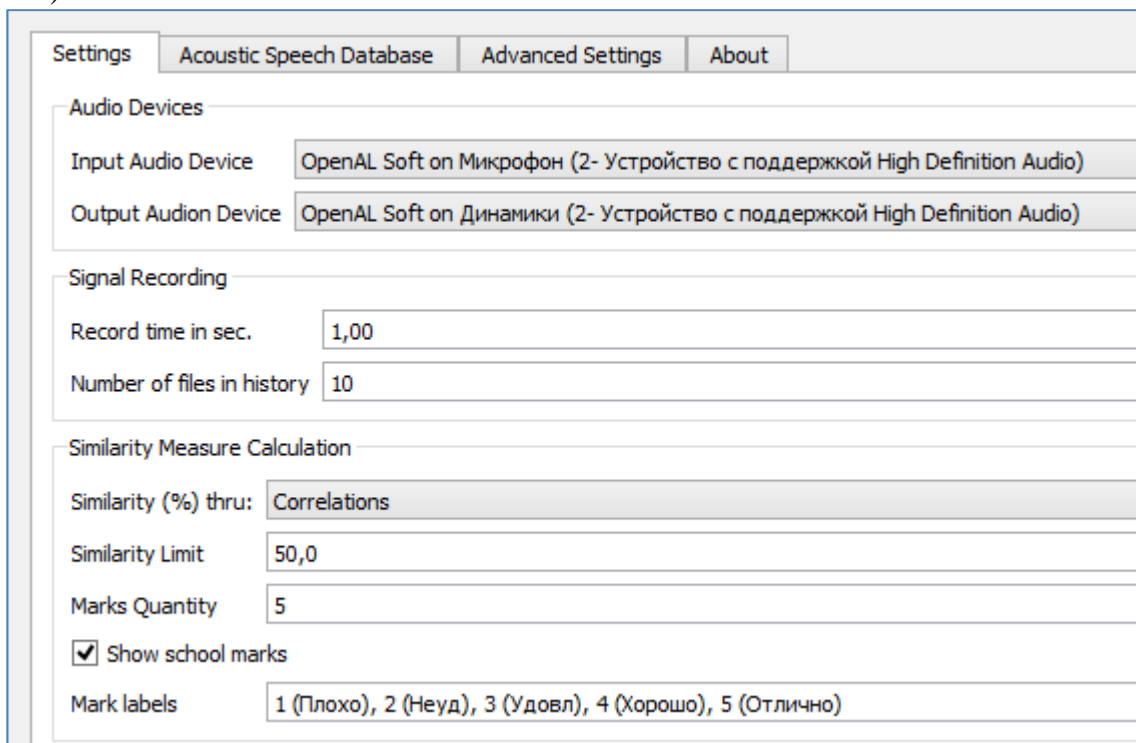


Рис. 3. Начальное окно настроек

В этом окне 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” указанного количества записываемых с микрофона фраз – **Number of files in history**,

В разделе **Similarity Measure Calculation** имеет возможность выбрать метод оценки сходства интонации произнесённой фразы с эталонной, определяемой одним из трёх способов. Посредством вычисления – **Similarity (%) thru**:

- 1) взаимной корреляции – **Correlations**,
- 2) среднего значения взаимного расстояния – **Average Distance**,
- 3) максимума локального расстояния - **Maximum Local Distance**,
- 4) среднего значения 3-х вышеуказанных оценок сходства – **Average**.

Выбранный метод вычисления интонационного сходства может быть использован затем при вычислении 5-ти бальной (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.

Две дополнительные настройки – **Acoustic Speech Database** и – **Advanced settings** - в главном окне (рис. 3) предназначены для настроек, осуществляемых разработчиками или «продвинутыми» пользователями программы (рекомендации по их использованию будут приведены ниже в разделах 5, 6). Кнопка – **About** открывает окно со сведениями о разработчиках и ряд других сведений.

3. Просмотр и прослушивание интонационных конструкций русской речи

Прокручивая страницу главного окна (рис. 2), пользователь получает возможность слухового и визуального ознакомления с основными интонационными конструкциями русской речи. Для этого, путём выбора с помощью курсора требуемой директории, например:

{1} **Примеры УМП – ИК1 «(ж) Он гуля+ет в парке с собакой»**

открывается окно, в котором отображается в графическом виде результаты интонационного анализа этой фразы (рис. 4)

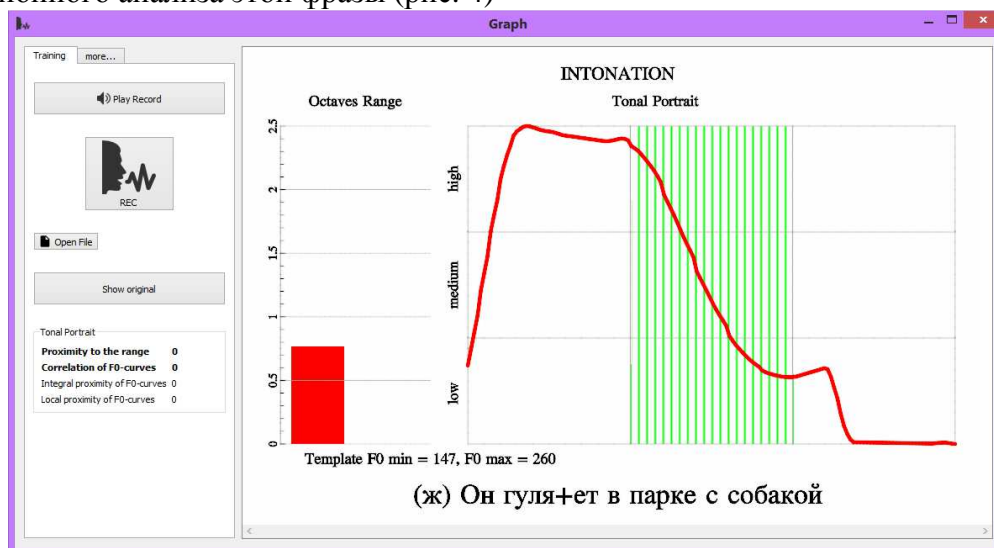


Рис. 4. Окно отображения результатов анализа

На рис. 4 красный столбец слева отображает диапазон изменения частоты основного тона, выраженный в октавах. **Oktaves Range** = (F0max / F0min) - 1.

Справа, красным цветом отображается линейный график УМП фразы, ядро которой помечено частыми вертикальными линиями. Внизу под графиками приведены минимальное и максимальное значения F0 для выбранной фразы, а также текст фразы, в котором ядерный гласный обозначен знаком «+».

Прослушивание выбранной фразы осуществляется путём нажатия кнопки «**Play Record**».

При нажатии кнопки «**Show original**» открывается дополнительное окно (рис. 5), в верхней части которого отображается осциллограмма сигнала фразы с указанием меток пред-ядра (красная линия), ядра (чёрная линия) и за-ядра (голубая линия). В средней части окна изображается кривая изменения F0 в реальном времени. На ней указаны участки пред-ядра, ядра и за-ядра, из которых формируется УМП, показанный на рис. 4. Функция «**Show original**» используется для контроля в тех редких случаях, когда возникают сомнения в правильности полученной УМП фразы.

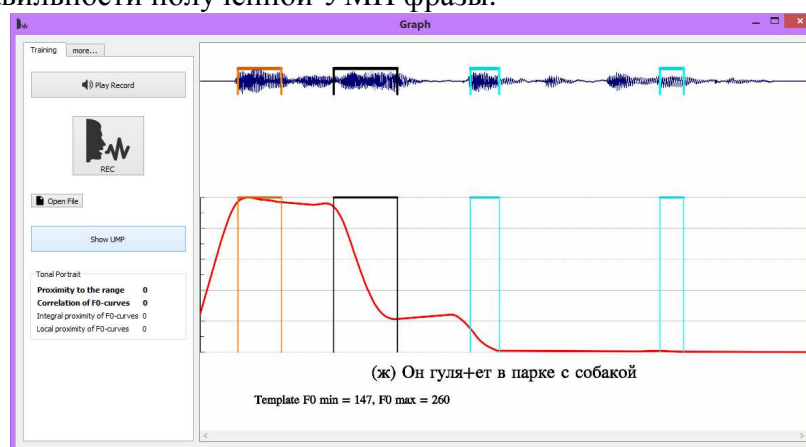


Рис. 5. Дополнительное окно отображения F0 в реальном времени

Листая страницу главного окна (рис. 2) сверху вниз, пользователю предоставляется возможность слухового и визуального ознакомления с **примерами основных интонационных конструкций** русской речи (ИК1 – ИК8), представленных в виде УМП, с **их попарным сопоставлением**, с **особенностями употребления ИК**, а также с реализацией ИК в диалогах, прозе, стихах и при пении.

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

При использовании данного **Приложения** в целях индивидуального интонационного тренинга при изучении русского как иностранного (РКИ), или для совершенствования имеющихся устно-речевых интонационных навыков, пользователь (после того, как он выбрал, просмотрел и прослушал требуемую фразу) может воспользоваться внешним или встроенным микрофоном. При этом он должен нажать большую кнопку «**Rec**» в центре левой секции окна (см. рис.6), дождаться короткого «бип-сигнала» и произнести в микрофон фразу, текст которой указан в нижней части окна на рис. 4. После записи и обработки программой введенного речевого сигнала пользователь услышит 2-й «бип-сигнал», а изображение в графическом окне (рис.4) заменится на изображение, показанное на рис. 6. В верхней части окна показаны результаты сравнения эталонной и произнесённой фраз: **Range (87%)** - близость по диапазону изменения F0 и **Shape (98%)** – близость по форме траектории F0. Рядом с

процентными оценками близости могут быть проставлены (если это необходимо) оценки по 5-ти бальной системе, как это показано на рис. 6.

На рис. 6 красный столбец слева отображает диапазон изменения F0 эталонной фразы, а коричневый – произнесённой. Справа, красным цветом отображается линейный график УМП эталонной фразы, а коричневый – произнесённой фразы. Внизу под графиками приведены минимальное и максимальное значения F0 эталонной и произнесённой фраз.

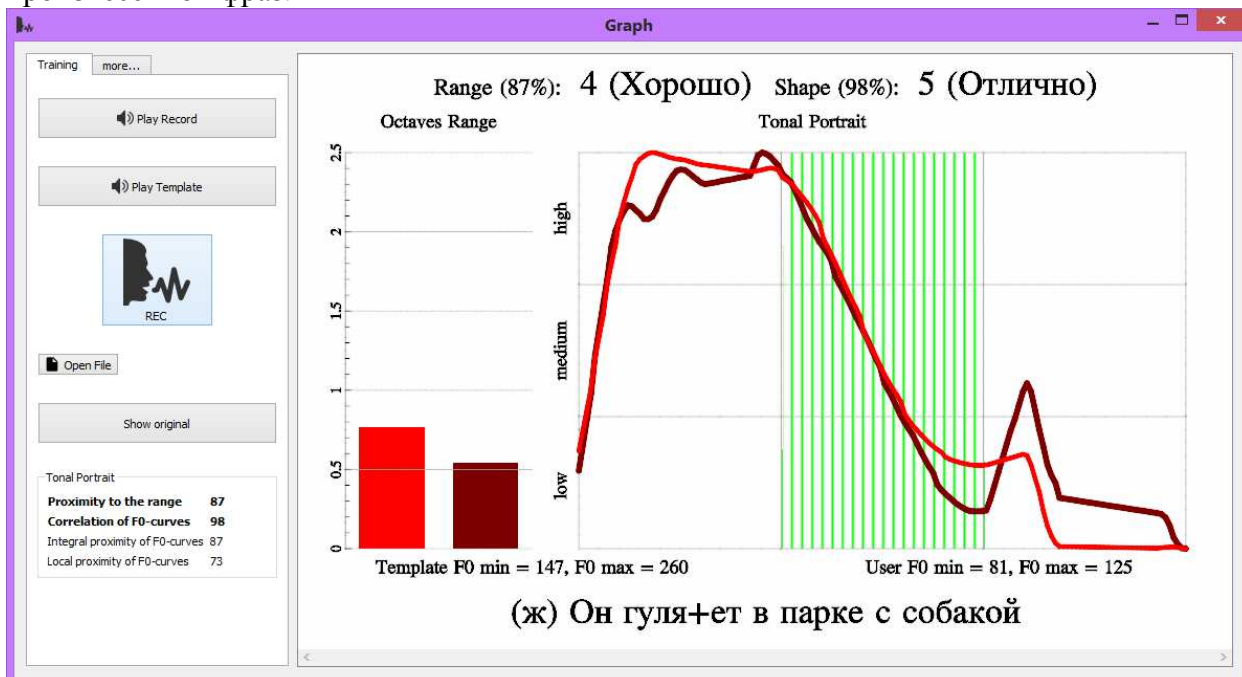


Рис. 6. Окно отображения результатов анализа и сравнения

Прослушивание выбранной эталонной фразы осуществляется путём нажатия кнопок «**Play Template**», а произнесённой фразы - «**Play Record**», находящихся в верхней части левой секции окна на рис.6.

При нажатии кнопки «**Show original**» открывается дополнительное окно (рис. 7), в верхней части которого отображаются осциллограммы сигналов обеих фраз 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. 6. 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.

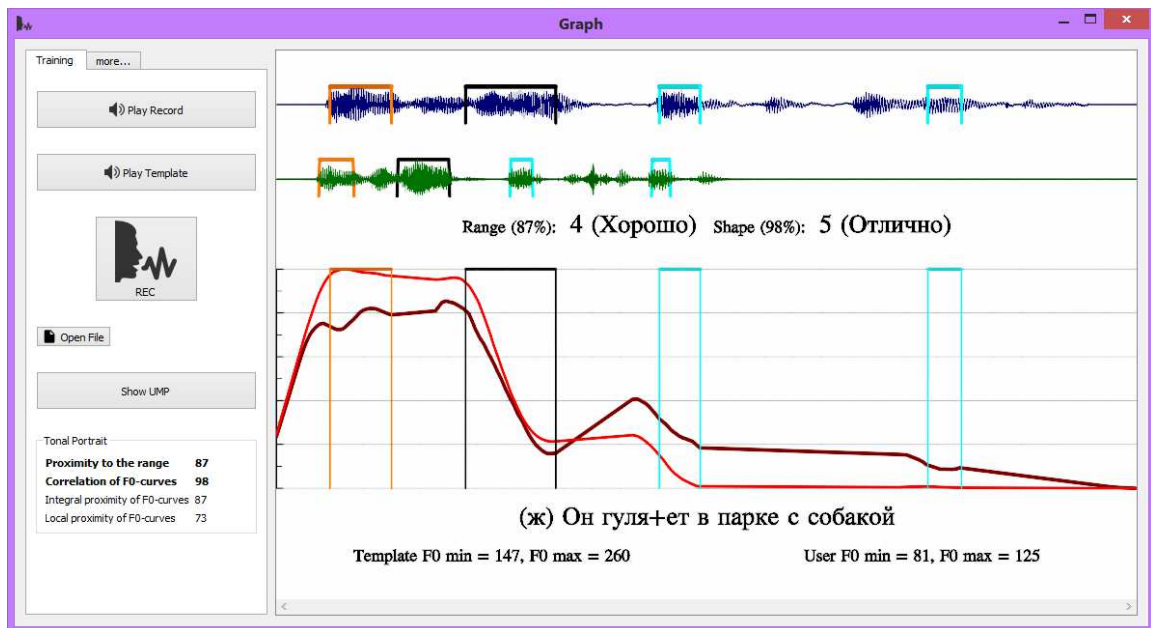


Рис. 7. Дополнительное окно отображения совмещённых траекторий F0

5. Сравнение интонации фраз от различных источников

Как уже указывалось, данное Приложение может быть использовано также в качестве инструмента в ряде научно-практических исследований. Например, Приложение может быть с успехом использовано в экспериментально-фонетических исследованиях, в ходе которых возникает необходимость сопоставления эталонной интонации с интонацией исследуемых фраз, полученных от различного рода источников (например, при сравнении интонации естественной и синтезированной речи). В этом случае, вместо использования внешнего или встроенного микрофона, используется кнопка “Open Instance File” в левой секции окна. При этом из специально созданной папки “TEST” выбирается речевой сигнал того же содержания, но от другого источника, в данном случае от синтезатора речи. Графический результат сравнения показан в правой части рис. 8.

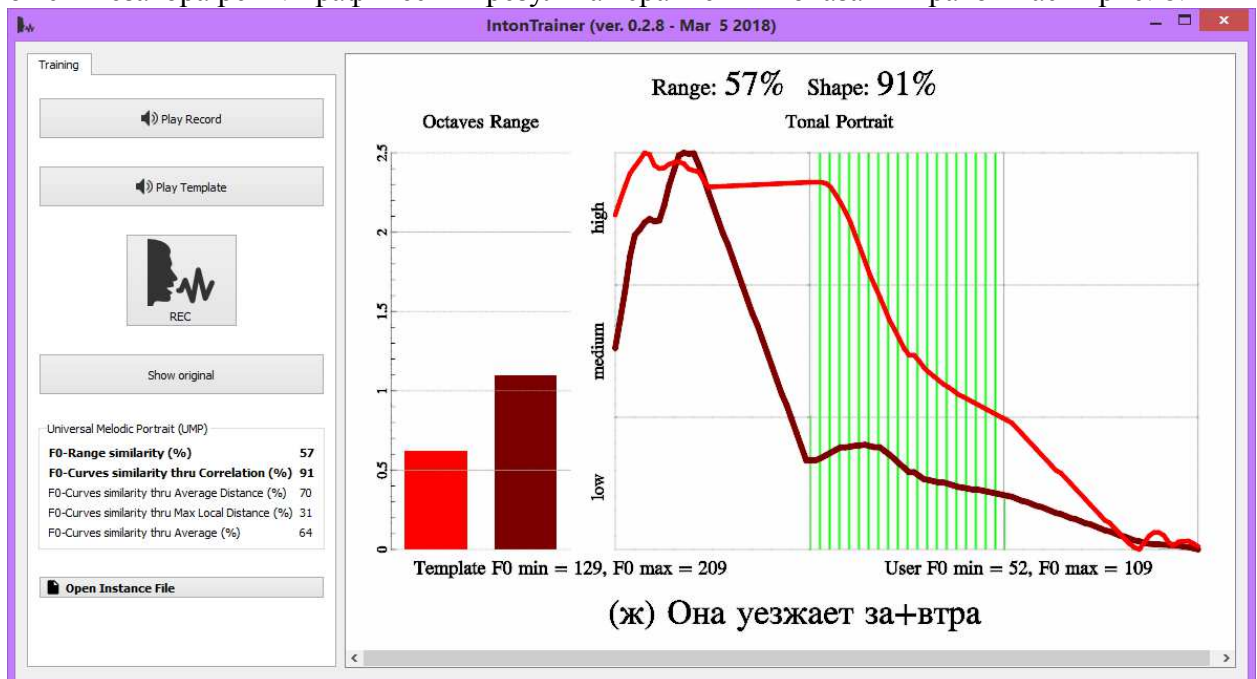


Рис. 8. Результат сравнения интонации естественной и синтезированной речи

В верхней части окна показаны результаты сравнения естественной и синтезированной фраз: **Range (57%)** - близость по диапазону изменения F0 и **Shape (91%)** – близость по форме траектории F0.

Более детально численные оценки сходства даны на рис.8 в левой части окна, как это представлено ниже:

F0-Range similarity (%)	57
F0-Curves similarity thru Correlation (%)	91
F0-Curves similarity thru Average Distance (%)	70
F0-Curves similarity thru Max Local Distance (%)	31
F0-Curves similarity thru Average (%)	64

6. Дополнительные настройки – «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» открывается окно (рис. 9), содержащее структурированный перечень эталонных фраз.

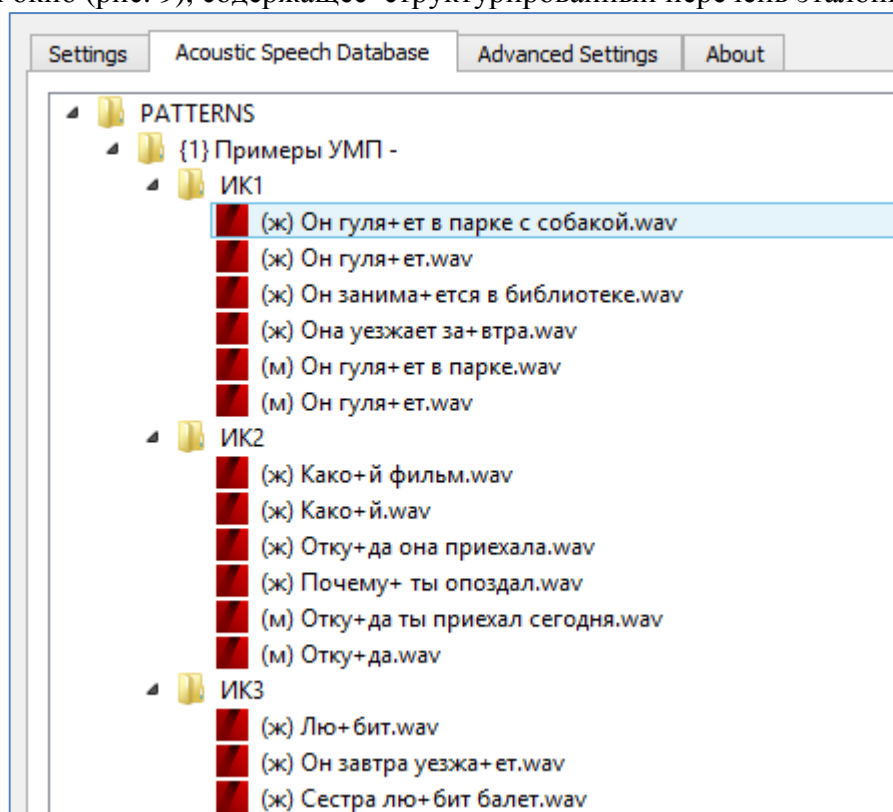


Рис. 9. Окно отображения структурированного перечня эталонных фраз

По двойному клику мыши на этом перечне открывается файл или окно проводника. При правом клике мыши по папке открывается выпадающее меню (рис. 10):

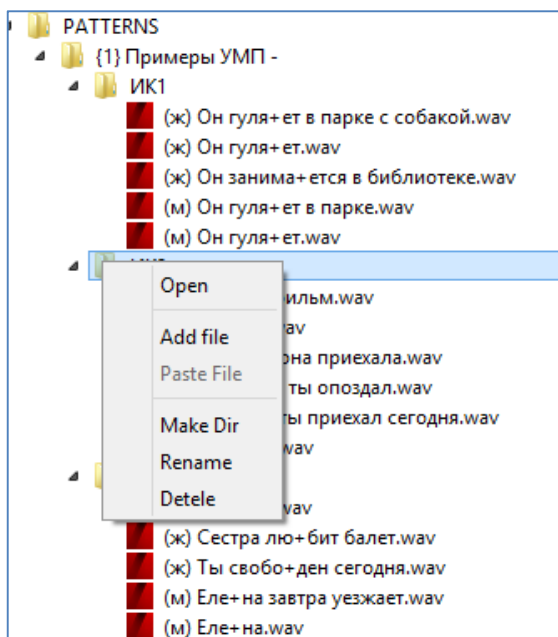


Рис. 10. Выпадающее меню при правом клике мыши по папке ИК2

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. 11):

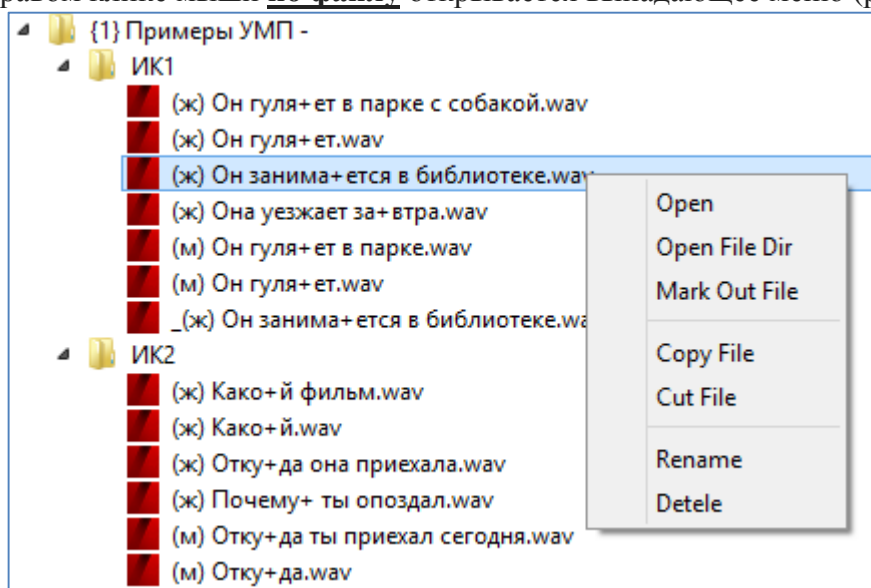


Рис. 11. Выпадающее меню при правом клике мыши по файлу: {(ж) Он занима+ется в библиотеке.wav}

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 (см. рис. 11).

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. 12 для одно-акцентной фразы: «**Он гуля+ет**», произнесённой женским голосом. Результат построения УМП этой фразы, полученный на основе такой разметки, представлен на рис. 13.

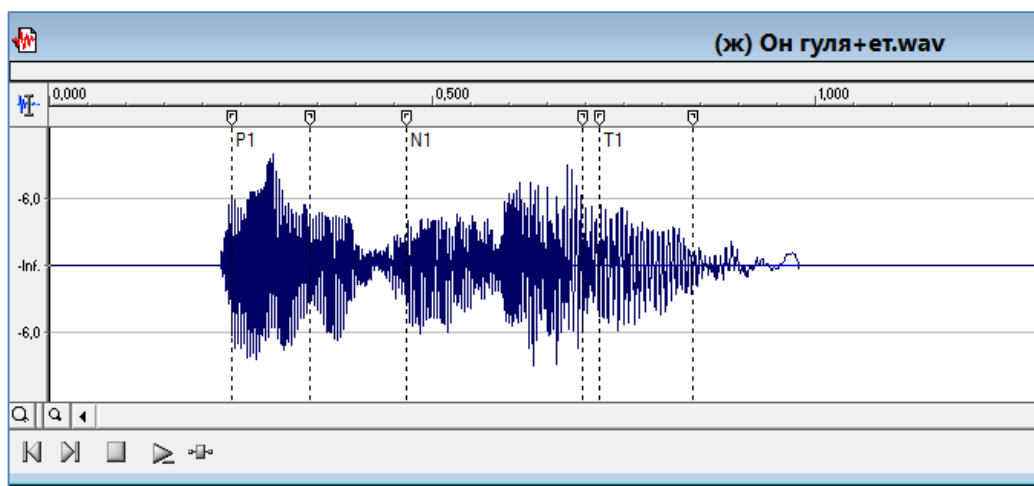


Рис. 12. Примеры разметки одно-акцентной эталонной фразы:
«(ж) Он гуля+ет»

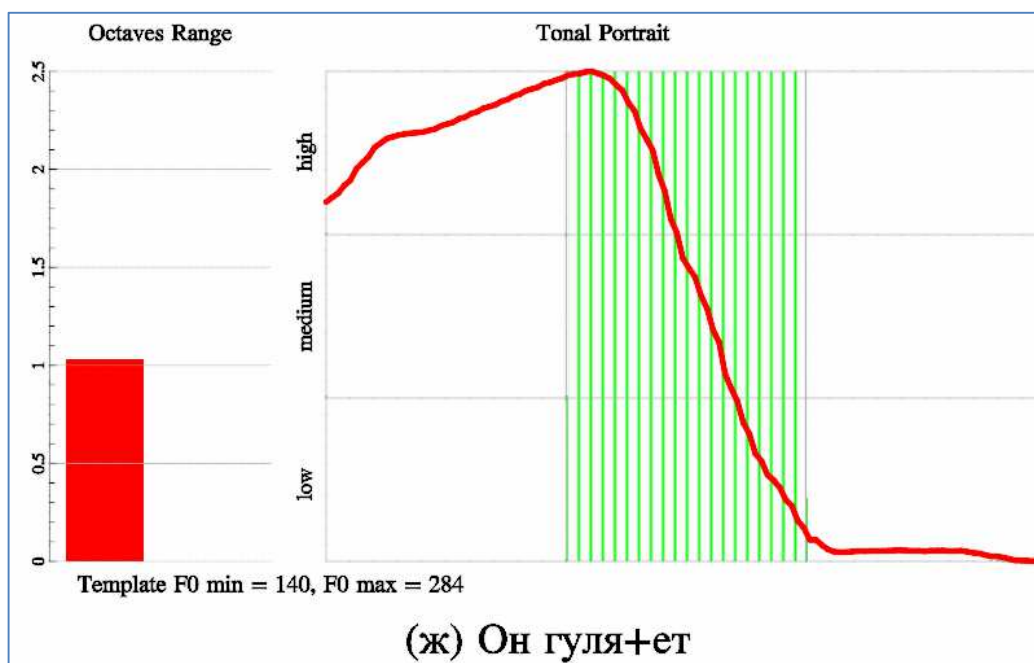


Рис. 13. Результат построения УМП фразы:
«(ж) Он гуля+ет»

На рис. 14 приведен пример разметки 2-х акцентной фразы: «(Ж) **Како+й сегодня де+нь!**», а на рис. 15 показан результат построения УМП этой фразы, полученный на основе такой разметки.

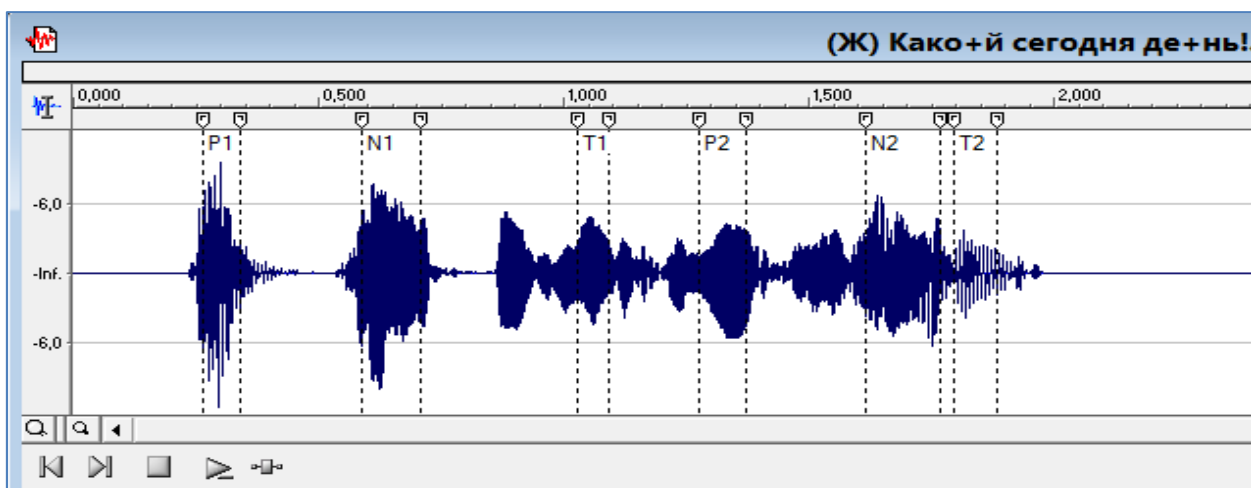


Рис. 14. Пример разметки 2-х акцентной фразы:
«(Ж) Како+й сегодня де+нь!».

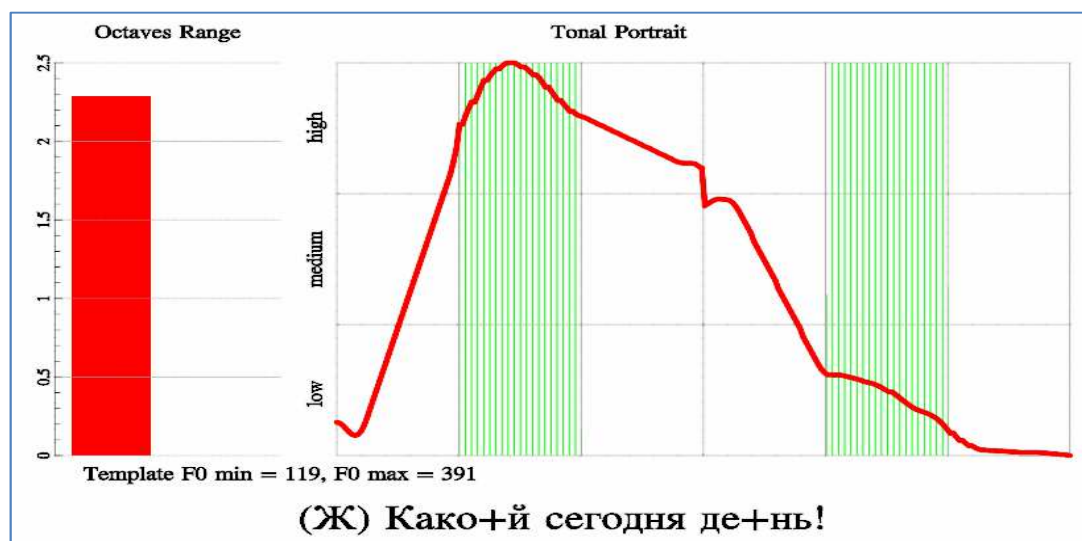


Рис. 15. Результат построения УМП фразы:
«(Ж) Како+й сегодня де+нь!».

Если, по каким-либо причинам определение положения элементов акцентной структуры фразы (P, N, T) является затруднительным, 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. Для примера на рис. 16 показан УМП фразы «Он гуля+ет в парке с собакой», построенный при наличии разметки на P, N, T – регионы, а на рис. 17 изображена траектория изменения F0 в том случае, когда каждому из голосовых участков фразы присвоен одинаковый индекс N. Создание такого файла было осуществлено путём выбора операции (**Mark Out File**) в разделе дополнительные настройки – «Acoustic Speech Database» (см. рис. 11).

Отметим, что автоматизация процесса разметки фраз на голосовые регионы постоянно совершенствуется. В конечном итоге предполагается полная автоматизация процесса, т.е. осуществление не только разметки сигнала фразы на голосовые регионы, но и автоматическое определение положения каждого из элементов акцентной структуры фразы: предъядра - P, ядра – N и зядра - T.

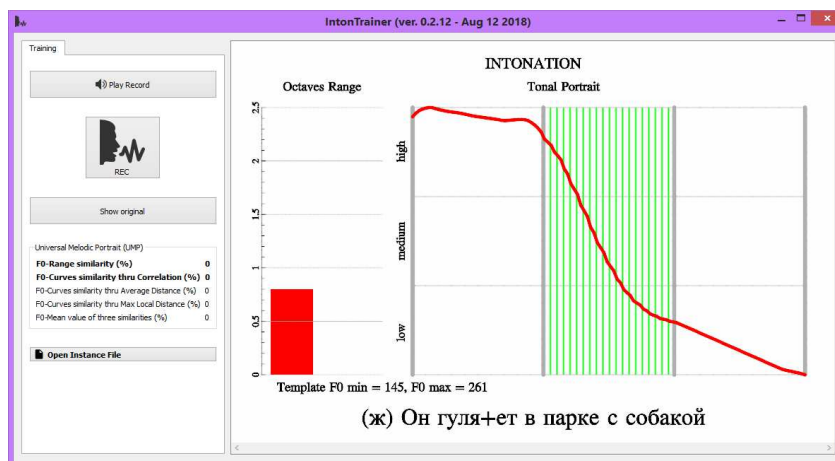


Рис. 16. Пример отображения при ручной разметке фразы на P, N, T – регионы

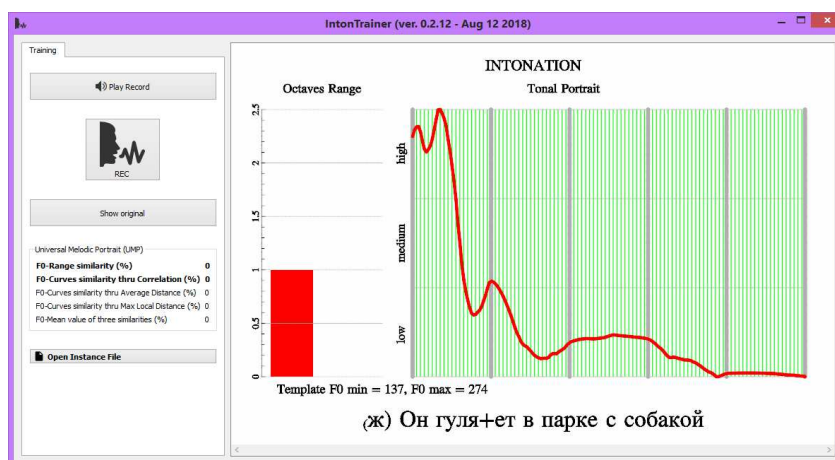


Рис. 17. Пример отображения при автоматической разметке фразы на голосовые участки, которым присвоен одинаковый индекс N

Чтобы узнать больше о теоретических основах разработки Приложения, см. дополнительную информацию, размещенную на этом сайте.

8. ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

На данном сайте пользователю предоставляется возможность бесплатного скачивания трёх демо версий интонационного тренажёра «IntonTrainer», предназначенных для обучения базовым интонационным конструкциям русского языка и 2-х вариантов английского языка (британского и американского). На сайте представлены также прототипы демо версий для немецкого и китайского языков. Кроме того, приведена демо версия с примерами обучения интонации для дикторов колл-центров на русском и английском языках, а также демо версия для обучения эмоциональной интонации на английском языке.

Как уже указывалось, программный комплекс «IntonTrainer» является открытой системой и пользователю предоставляются широкие возможности для корректировки, обновления и создания эталонных баз данных для различных языков и разнообразных практических приложений. При этом мы рекомендуем использование варианта автоматической разметки эталонных звуковых файлов путём выбора операции (**Mark Out File**) в разделе дополнительные настройки – «Acoustic Speech Database».

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