

Руководство пользователя
Vital Voice Tools

Система синтеза русской речи по тексту
TTS Vital Voice

(С) ООО ЦРТ, 2010-2015
Все права защищены

СОДЕРЖАНИЕ

<u>1. ПРИЛОЖЕНИЕ VITAL VOICE TOOLS.....</u>	<u>3</u>
<u>2. СИНТЕЗ.....</u>	<u>5</u>
<u>3. ПОЛЬЗОВАТЕЛЬСКИЙ СЛОВАРЬ.....</u>	<u>16</u>
<u>4. ПРОИЗНОШЕНИЕ ФРАЗ.....</u>	<u>29</u>
<u>5. ПРА-ТРАНСКРИПЦИЯ.....</u>	<u>52</u>

1. ПРИЛОЖЕНИЕ VITAL VOICE TOOLS

Программа **Vital Voice Tools** (далее VVT) это графический интерфейс, предназначенный для использования и настройки процессора Vital Voice Text-To-Speech (далее TTS).

VVT позволяет синтезировать произвольный текст, сформировать и редактировать пользовательский словарь, настроить произношение и задать IPA-транскрипцию для заданных фраз и слов.

Внешний вид основного окна приложения показан на *рис. 1*.

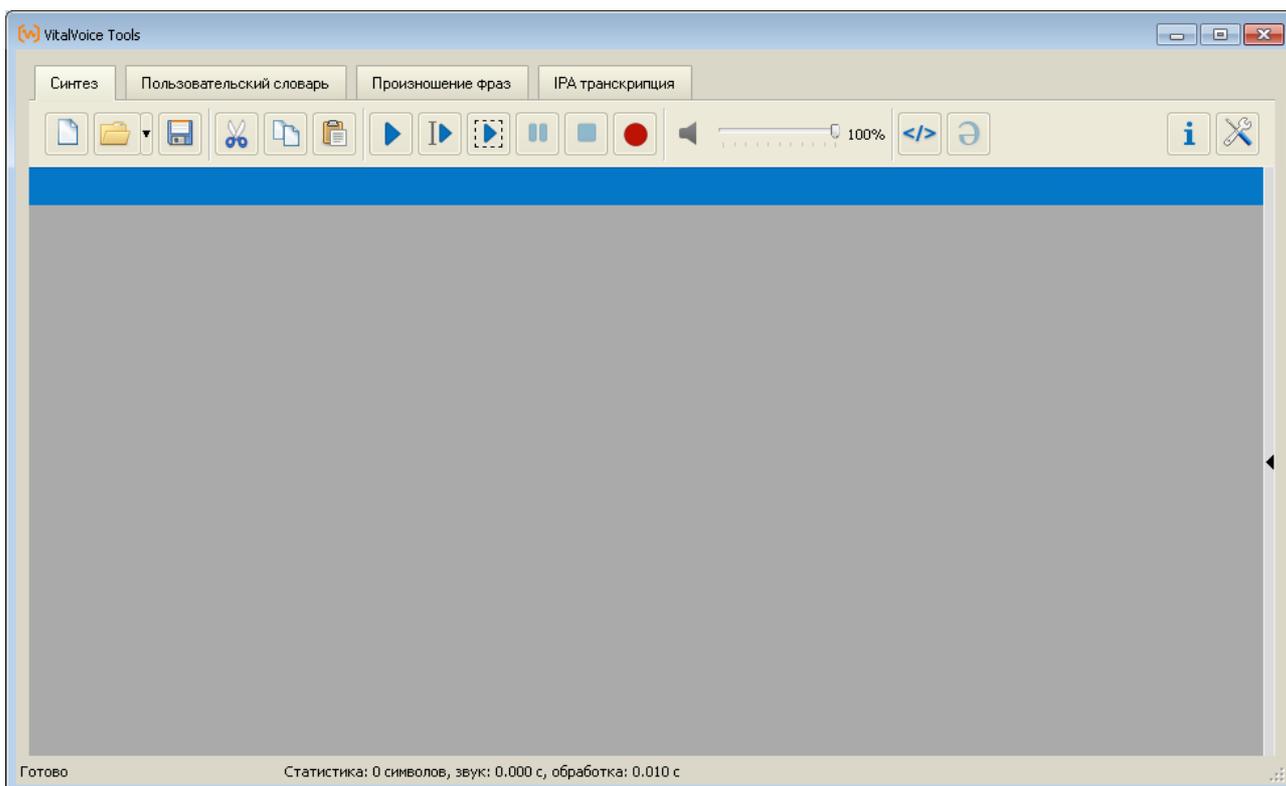


Рис. 1. Графический интерфейс приложения.

1.1. Состав приложения

Приложение VVT состоит из четырех основных частей (модулей):

- Модуля синтезатора, позволяющего воспроизводить произвольный текстовый файл путем произнесения его содержимого при помощи установленного процессора Vital Voice TTS - движка синтеза.
- Модуля редактирования пользовательского словаря, позволяющего вносить изменения в словарь пользователя и добавлять в него незнакомые слова, встретившиеся в произносимом тексте.

- Модуля настройки пользовательских произношений, позволяющего пользователю настраивать произношения отдельных фраз путем модификации транскрипции и управления процедурой подбора звуковых единиц.
- Модуля редактирования словаря IPA-транскрипций, позволяющего для любого слова задать его транскрипцию, записав ее в виде последовательности символов международного фонетического алфавита IPA.

Каждый модуль представлен в виде отдельной вкладки на основном окне приложения.

1.2. Установка приложения

Для корректной работы приложения на компьютере пользователя должны быть установлены следующие компоненты:

- процессор Vital Voice TTS;
- одна или несколько голосовых баз;
- .NET Framework;

1.3. Запуск приложения

Приложение запускается без параметров. При запуске происходит загрузка модуля «Синтез» на первой вкладке основного окна приложения. Загрузка остальных модулей происходит при первом переключении на вкладку соответствующего модуля.

1.4. ini файл

Состояние интерфейса приложения (положение и размер окна, положение разделительных полос, ширина столбцов таблиц, выбранный голос и т. п.) сохраняется в ini файле в папке

`ApplicationData\Speech Technology Center\VVTools`

где

`ApplicationData = C:\Documents and Settings\<user_name>\Application Data`

(указано обычное расположение этой папки для Windows XP).

2. СИНТЕЗ

Модуль синтезатора (Синтез) представляет собой графический интерфейс для воспроизведения текстов. Синтез позволяет воспроизводить произвольный текстовый файл путем произнесения его содержимого при помощи установленного движка синтеза.

Функциональность Синтеза:

- Открытие, отображение и редактирование текста в формате ТХТ
- Сохранение текста в формате ТХТ с возможностью выбора имени и пути сохранения файла (по умолчанию при сохранении указывается последний путь сохранения файла)
- Воспроизведение текста с возможностью предварительного выбора голоса
- Возможность установки громкости воспроизведения синтеза в диапазоне от 0% до 100%
- Изменение темпа звучания от 0.5 до 4.0
- Изменение тона звучания от 0.5 до 2.0
- Изменение тембра звучания от -1 до +1
- Изменение длительности пауз от 0.1 до 5.0
- Изменение частоты дискретизации сигнала от 8000 до 44100 Гц
- Поддержка тегов SSML
- Поддержка алфавита IPA
- Режим "караоке"
- Управление синтезом: воспроизведение, воспроизведение от курсора, воспроизведение выделенного фрагмента, остановка воспроизведения, пауза.
- Сохранение синтезированного звука в файл без воспроизведения с выбором имени файла, пути сохранения, формата (MP3 32-160kbps или WAV).
- Пакетный режим синтеза текстовых файлов с сохранением синтезированного звука в WAV или MP3 файл.

2.1. Интерфейс и использование модуля

Графический интерфейс модуля «Синтез» представлен на *рис. 2.1*.

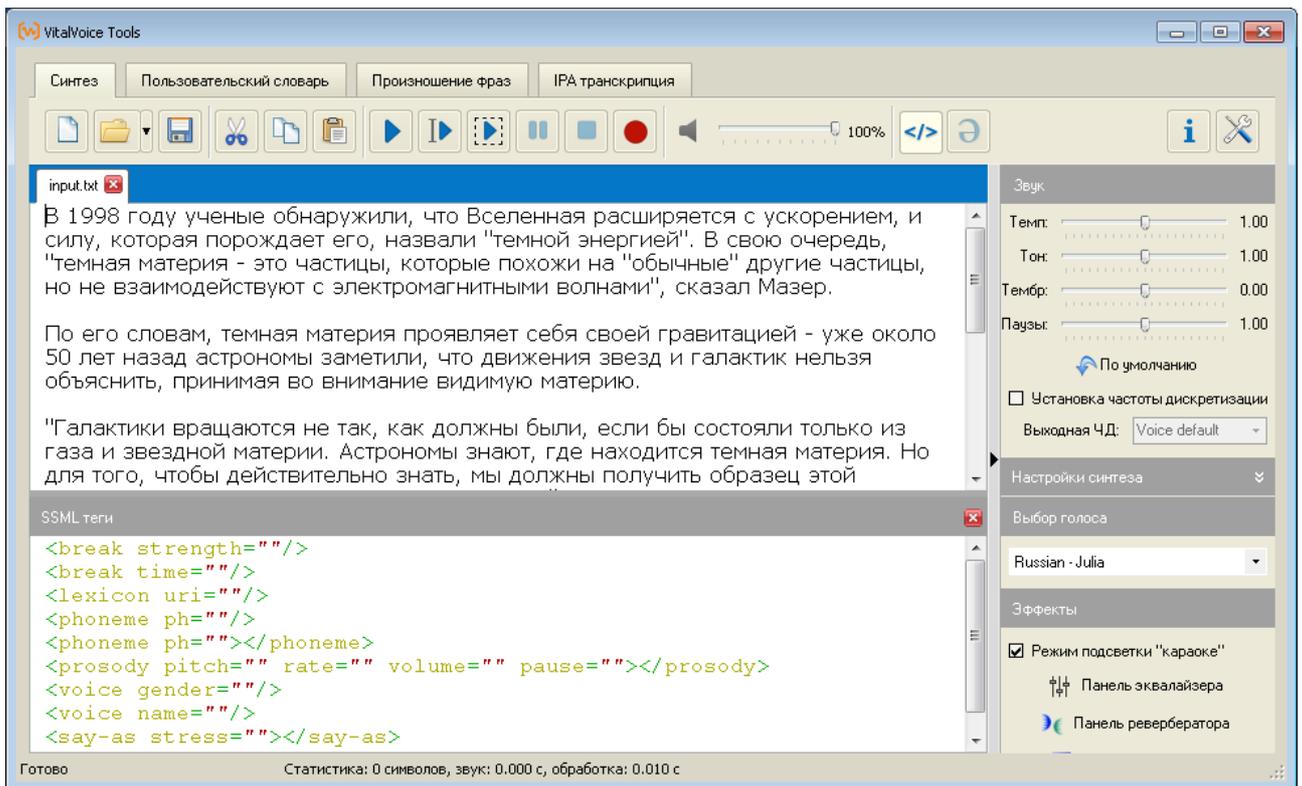


Рис. 2.1. Графический интерфейс модуля «Синтез».

Графический интерфейс модуля предоставляет собой:

- окно отображения текста,
- панель инструментов с кнопками,
- нижнюю панель, в которой отображается окно ввода SSML тегов и виртуальная клавиатура IPA,
- боковую панель, содержащую элементы управления звучанием.

2.1.1. Окно отображения текста

Окно предназначено для отображения текста, вводимого пользователем или загружаемого из файла. Непосредственно над окном отображения текста находится панель документов, каждая вкладка которой соответствует открытому или созданному пользователем документу. Окно может одновременно отображать только один текстовый документ, при этом выбор отображаемого документа производится путем выбора соответствующей вкладки на панели документов.

Текст, отображаемый в окне, может быть воспроизведен путем нажатия кнопок на панели инструментов (см. разд. 2.1.2. *Панель инструментов*).

2.1.2. Панель инструментов

Панель инструментов содержит кнопки управления документами (создать документ, открыть документ, сохранить документ), кнопки редактирования

текста (вырезать, копировать, вставить), кнопки управления воспроизведением текста (произнести весь текст, произнести начиная от курсора, произнести выделенный фрагмент, пауза, остановить), регулятор громкости воспроизведения, кнопки управления нижней панелью (см. разд. 2.1.3. *Нижняя панель*), а также кнопку отображения информации о программе (см. разд. 2.1.5. *Информация о программе*) и кнопку вызова диалогового окна конфигурации модуля (см. разд. 2.1.6. *Диалог конфигурации модуля*).

Кнопки управления документами и редактирования текста работают обычным образом и дополнительных комментариев не требуют. Следует отметить особенность работы кнопки сохранения документа: всякий раз при сохранении документа требуется указать в появляющемся диалоге имя файла – кнопка работает по принципу «сохранить как». Нажатие на «стрелку вниз», находящуюся справа от кнопки открытия документа, открывает список из 10 последних открытых или сохраненных документов; щелчок по названию документа из этого списка открывает выбранный документ.

При нажатии кнопки произнесения текста (воспроизведение, Play) синтезатор речи начинает произносить текст, отображаемый в окне. При нажатии кнопки «произнести от курсора» или «произнести выделенный текст» происходит воспроизведение соответствующей части текста.

Кнопка «Стоп» прерывает воспроизведение текста, при последующем нажатии кнопки «Play» текст будет воспроизводиться от начала.

Кнопка «Пауза» приостанавливает воспроизведение текста. Повторное нажатие кнопки «Пауза» возобновляет воспроизведение текста с того места, на котором оно было остановлено.

Кроме воспроизведения текста из открытого документа приложение позволяет воспроизводить текст из буфера обмена Windows. Для этого следует скопировать текст в буфер обмена и нажать горячую клавишу Ctrl-F11. Будет создан виртуальный документ «*/clipboard», в него будет скопирован текст из буфера обмена, после чего будет воспроизведен текст из этого документа, причем текущий процесс воспроизведения, если таковой продолжался, будет предварительно остановлен.

В процессе воспроизведения текста в основном окне выделяется слово, произносимое в данный момент (режим «караоке»). Этот режим отключается соответствующей галочкой на боковой панели (см. разд. 2.1.4. *Боковая панель*).

Регулятор громкости производит установку громкости воспроизведения произносимого текста. Громкость регулируется независимо от системного микшера. Если при установке регулятора громкости на максимум громкость воспроизведения недостаточна, следует проверить настройки системного микшера.

Кнопки управления нижней панелью («Добавление SSML тегов» и «Виртуальная клавиатура IPA») представляют собой кнопки с фиксацией («залипающие»), при нажатии которых отображаются соответственно панель ввода SSML тегов и виртуальная клавиатура IPA (см. разд. 2.1.3. *Нижняя панель*).

Кнопка отображения информации о программе вызывает окно информации о программе (см. разд. 2.1.5. *Информация о программе*), в котором представлена информация о программе и некоторая статистика.

Кнопка «Настройки» вызывает диалоговое окно конфигурации модуля (см. разд. 2.1.6. *Диалог конфигурации модуля*), содержащего некоторые дополнительные настраиваемые параметры.

2.1.3. Нижняя панель

В нижней панели могут отображаться либо панель ввода SSML тегов, либо виртуальная клавиатура IPA, в зависимости от положения текстового курсора в окне отображения текста.

При нахождении курсора где-либо в тексте вне SSML тегов становится доступна панель ввода SSML тегов, при этом кнопка «Добавление SSML тегов» активна и управляет отображением этой панели. При нажатии этой кнопки панель показывается, при отпускании – скрывается.

По двойному щелчку на какой-либо строке в панели ввода SSML тегов соответствующий SSML тег добавляется в окно ввода текста в позицию курсора. При этом если в окне был выделен текст, то он либо замещается, либо обрамляется вводимым тегом. Замещение происходит в том случае, если тег закрытый, например `<break time="" />`, в случае открытого тега выделенный текст вставляется между открывающим и закрывающим тегами, например `<phoneme ph="">выделенный текст</phoneme>`.

При нахождении текстового курсора в основном окне внутри уже введенного SSML тега панель ввода SSML тегов скрывается, так как новый тег не может быть введен внутрь уже имеющегося тега.

При нахождении текстового курсора в основном окне внутри атрибута `ph` тега `phoneme` (`<phoneme ph=""|">`) становится доступна виртуальная клавиатура IPA (рис. 2.1.3).

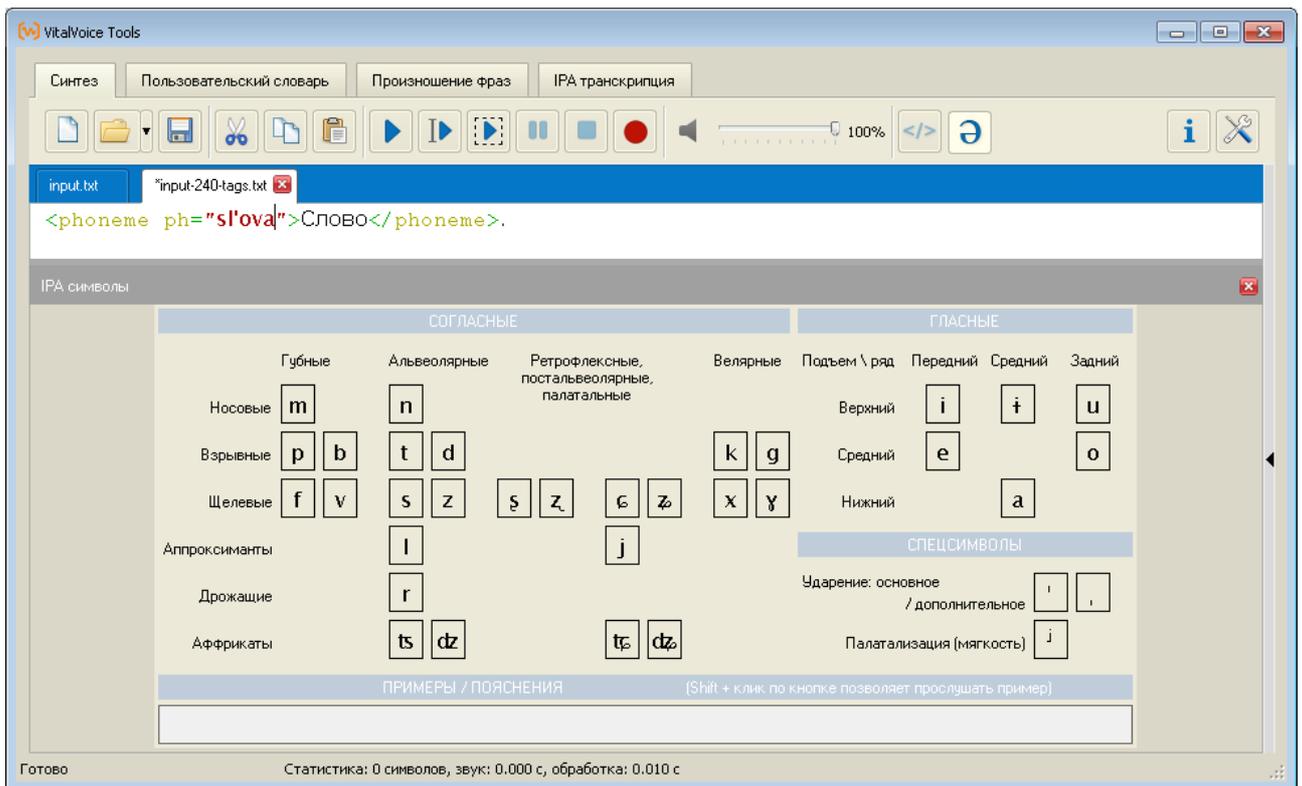


Рис. 2.1.3. Виртуальная клавиатура IPA.

При этом кнопка «Виртуальная клавиатура IPA» активна и управляет отображением этой клавиатуры. При нажатии этой кнопки клавиатура показывается, при отпускании – скрывается.

Если нижняя панель (панель ввода SSML тегов или виртуальная клавиатура IPA) была открыта пользователем, то при перемещении текстового курсора внутри текста управление отображением панели происходит автоматически.

2.1.4. Боковая панель

Боковая панель содержит регуляторы темпа и тона произнесения текста, тембра голоса диктора, длительности пауз, элементы управления преобразованием частоты дискретизации выходного звукового сигнала, раскрываемую панель настроек синтеза, выпадающий список выбора голоса синтеза, выключатель режима «караоке», а также кнопки вызова панелей эквалайзера и ревербератора.

Регулятор темпа устанавливает быстроту произнесения текста относительно естественной для выбранного диктора (голоса). Естественная быстрота произнесения соответствует значению темпа 1.00. Доступно регулирование в пределах от 0.5 (медленно) до 4.0 (очень быстро).

Регулятор тона устанавливает частоту основного тона произносимого текста относительно естественной для выбранного диктора (голоса).

Естественная частота основного тона соответствует значению тона 1.00. Доступно регулирование в пределах от 0.5 (вдвое ниже) до 2.0 (вдвое выше).

Регулятор тембра изменяет тембр голоса для выбранного диктора. Естественный тембр соответствует значению 0.0. Доступно регулирование в пределах от -1.0 («монстр») до +1.0 («Буратино»).

Регулятор длительности пауз устанавливает длительность пауз при произнесении текста относительно естественной для выбранного диктора (голоса). Этот параметр также влияет на быстроту произнесения текста. Естественная длительность пауз соответствует значению 1.00. Доступно регулирование в пределах от 0.1 (очень короткие) до 5.0 (очень длинные).

Кнопка «По умолчанию» восстанавливает значения по умолчанию для темпа, тембра, основного тона и длительности пауз.

Левый щелчок по названию любого из регуляторов на боковой панели восстанавливает для данного регулятора значение по умолчанию.

Галочка «ресэмплинг» включает преобразование частоты дискретизации выходного сигнала к желаемому значению. Значение выбирается в выпадающем списке и может быть задано любым в диапазоне от 8 до 44.1 кГц, либо может быть установлено в значение по умолчанию для данного голоса, как если бы галочка «ресэмплинг» была выключена.

Раскрываемая панель настроек синтеза позволяет управлять параметрами расстановки пауз и обработки переводов строки. Функциональность элементов управления понятна из их названия. Следует отметить, что границы параграфов влияют на длительность пауз: пауза между параграфами более длительна, чем пауза между предложениями.

Выпадающий список выбора голоса синтеза позволяет выбрать голос (диктора), которым будет синтезироваться произносимый текст. Количество голосов в списке зависит от числа установленных в системе голосов для синтеза.

Галочка «режим подсветки «караоке»» включает выделение произносимого в данный момент слова в окне отображения текста.

Кнопки «панель эквалайзера» и «панель ревербератора» вызывают соответственно панель управления параметрами эквалайзера (см. разд. 2.1.6.) и панель управления параметрами реверберации (см. разд. 2.1.8. *Параметры реверберации*).

Кнопка «Пакетная обработка» вызывает диалоговое окно пакетной обработки файлов (см. разд. 2.1.9. *Пакетная обработка файлов*).

Боковая панель может быть скрыта нажатием кнопки, расположенной непосредственно слева от этой панели.

2.1.5. Информация о программе

Окно информации о программе (рис. 2.1.5) вызывается путем нажатия кнопки «i» на панели инструментов.

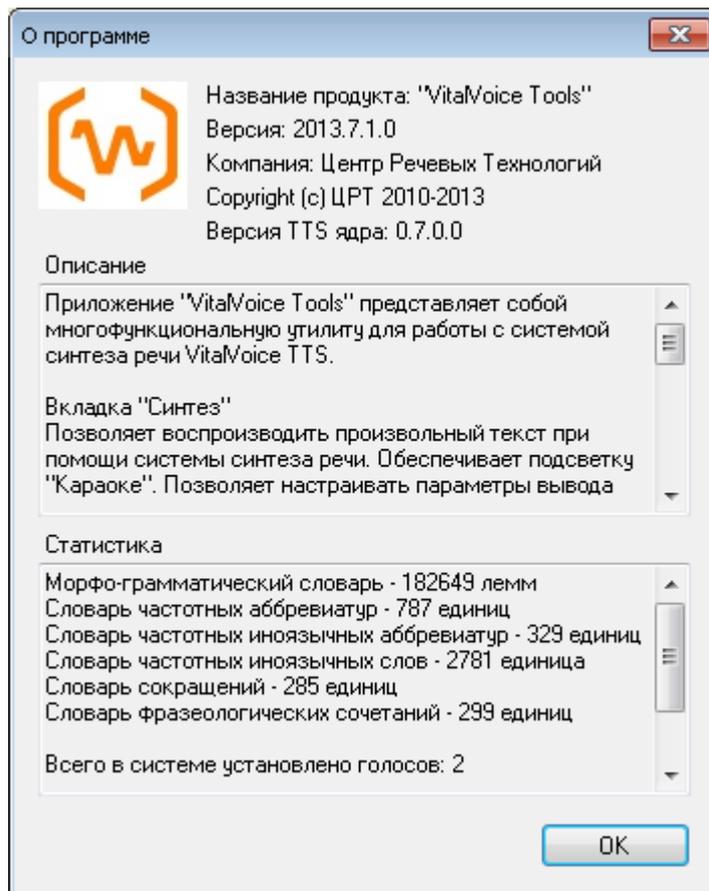


Рис. 2.1.5. Информация по программе.

В этом окне представлена информация о программе: название продукта, версия, производитель и краткое описание, а также некоторая статистика по используемым словарям и голосовым базам.

2.1.6. Диалог конфигурации модуля

Окно диалога конфигурации модуля (рис. 2.1.6) вызывается путем нажатия кнопки «Настройка» на панели инструментов.

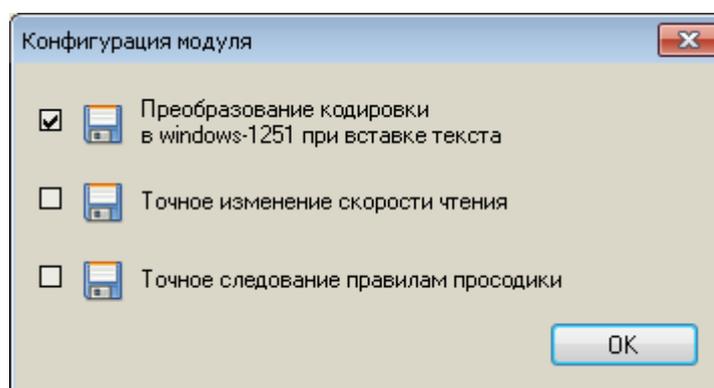


Рис. 2.1.6. Диалог конфигурации модуля.

Данное окно конфигурации предназначено для изменения параметров конфигурации модуля. Возможно изменение следующих параметров.

Галочка «Преобразование кодировки в windows-1251 при вставке текста». При включенной галочке перекодировка выполняется, при отключенной – нет. Перекодировку следует включать для русского языка при копировании текста из текстовых редакторов типа MS Word во избежание появления в тексте невидимых символов, таких как символ переноса, которые могут негативно влиять на синтез речи, вызывая нежелательные паузы. Отключение перекодировки имеет смысл для сохранения в копируемом тексте диакритик и прочих символов Unicode, например, для английского языка.

Галочка «Точное изменение скорости чтения». При включении галочки темп речи изменяется посредством модуля модификации звука; в этом режиме производится точная модификация, но возможно появление роботизированного звучания. При отключении галочки темп речи изменяется посредством изменения параметров интонационной модели; в этом режиме естественность звучания сохраняется.

Галочка «Точное следование правилам просодики». При включении задействуется особый режим полной модификации звука по всем параметрам с точной привязкой к интонационному контуру.

2.1.7. Трехполосный эквалайзер

В окне эквалайзера (рис. 2.1.7) пользователь может управлять параметрами эквалайзера: усилением по частотам и границами частотных полос.



Рис. 2.1.7. Трехполосный эквалайзер.

Эквалайзер можно совсем отключить, сняв галочку «Включить» на панели управления. Соответствующей кнопкой можно установить значения по умолчанию: усиление = 0, границы частотных полос = 300 и 3000 Гц.

2.1.8. Параметры реверберации

В окне управления параметрами реверберации (рис. 2.1.7) пользователь может управлять параметрами реверберации, либо выбирая одну из предлагаемых настроек, либо вручную регулируя задержку сигнала «эхо».

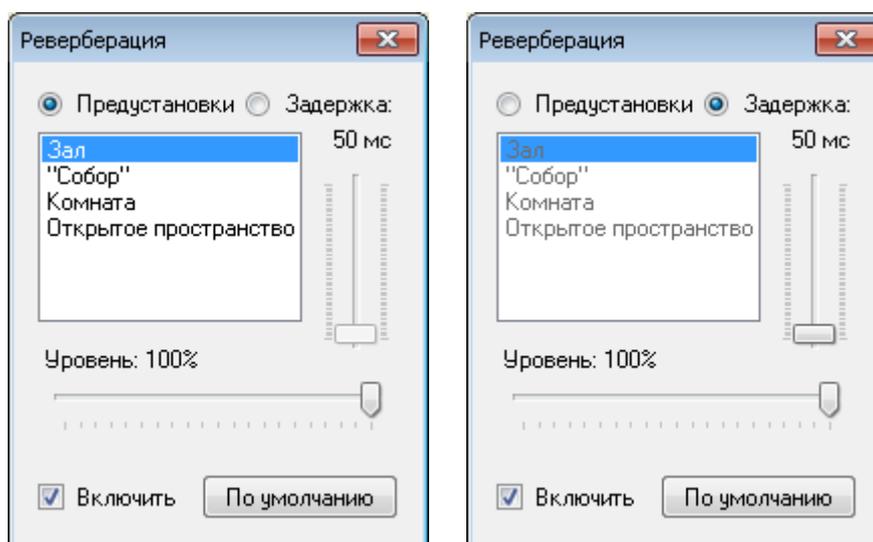


Рис. 2.1.8. Параметры реверберации.

В обоих случаях пользователь может управлять уровнем отраженного сигнала, используя соответствующий регулятор.

Реверберацию можно совсем отключить, сняв галочку «Включить» на панели управления. Соответствующей кнопкой можно установить значения по умолчанию: использование первой предустановки из списка с уровнем 100%.

2.1.9. Пакетная обработка файлов

Окно пакетной обработки файлов (рис. 2.1.9.1) вызывается путем нажатия на кнопку «Пакетная обработка», расположенную на боковой панели.

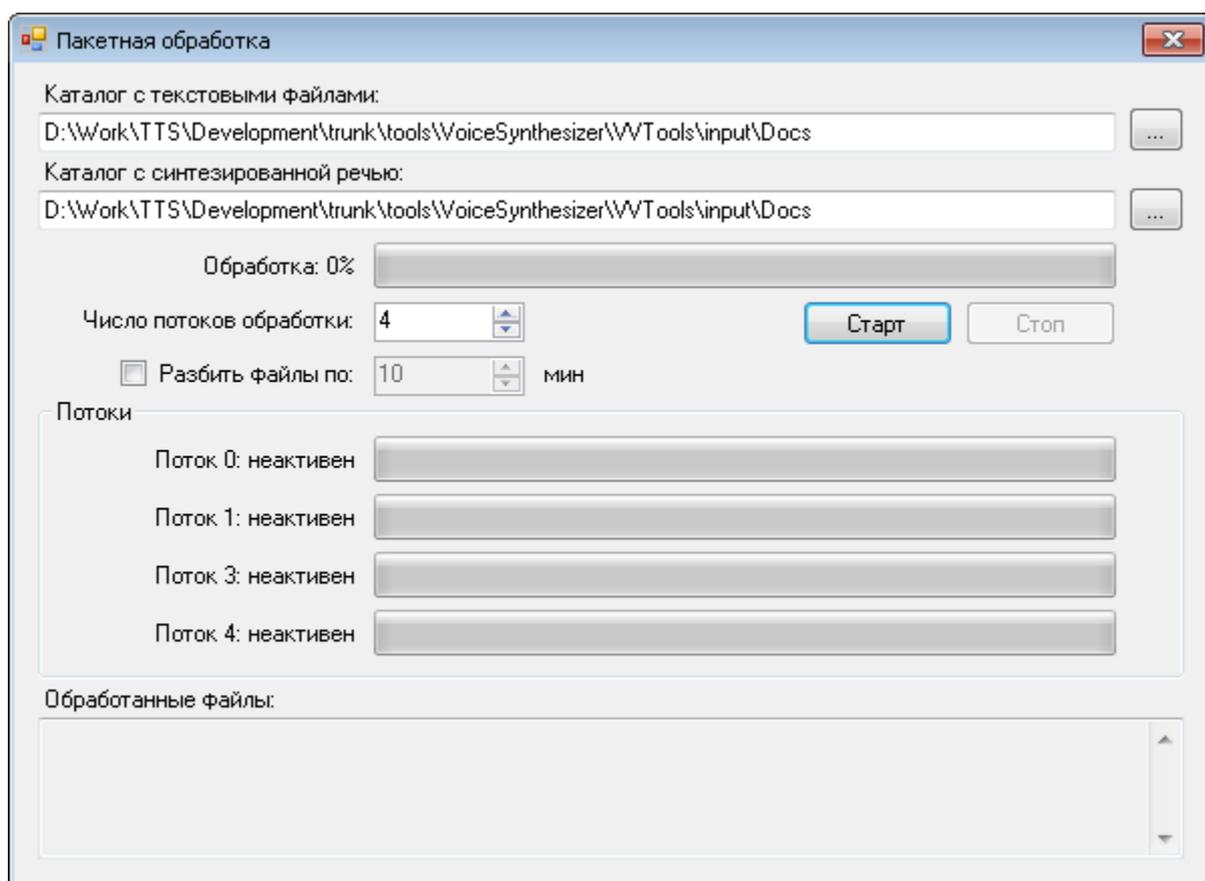


Рис. 2.1.9.1. Диалог пакетной обработки файлов.

При помощи данного диалога пользователь может синтезировать звук из нескольких текстовых файлов, при этом для каждого обработанного текстового файла будет создан и сохранен соответствующий звуковой файл.

Исходные текстовые файлы, предназначенные для обработки, следует скопировать в отдельную папку и выбрать эту папку в диалоговом окне, нажав кнопку «...» справа от строки «Каталог с текстовыми файлами». При этом путь к выбранной папке будет отображаться в указанной строке.

Папка, в которую будут сохраняться создаваемые звуковые файлы, указана в строке «Каталог с синтезированной речью». Эту папку можно выбрать, нажав кнопку «...» справа от указанной строки.

В окошке «число потоков обработки» можно указать число потоков, используемых для синтеза. Допустимые значения от 1 до 16, рекомендуется задавать это число равным числу процессорных ядер, имеющихся в системе.

Пара элементов управления «Разбивать файлы по N минут» позволяет синтезировать звуковые файлы с соответствующим ограничением по длительности. При превышении длительности файла создается следующий файл (file.wav, file0001.wav, file0002.wav и т.д.) Переход к очередному файлу выполняется после завершения синтеза текущего предложения.

После того, как будут выбраны папка, содержащая текстовые файлы, и папка, в которую будут сохраняться создаваемые звуковые файлы, можно запустить пакетную обработку, нажав кнопку «Старт».

Ход процесса обработки файлов будет отображаться на индикаторах прогресса, а также в окошке «Обработанные файлы». Общий индикатор прогресса «Обработка» показывает отношение количества уже обработанных файлов к общему количеству файлов, предназначенных для обработки. Индикаторы прогресса каждого потока отражают процесс обработки файла, синтезируемого в соответствующем потоке. В окошко «Обработанные файлы» заносятся уже обработанные файлы.

По окончании обработки в окошко «Обработанные файлы» добавляется слово «ГОТОВО» для сообщения о завершении процесса пакетной обработки и готовности обработанных файлов (рис. 2.1.9.2).

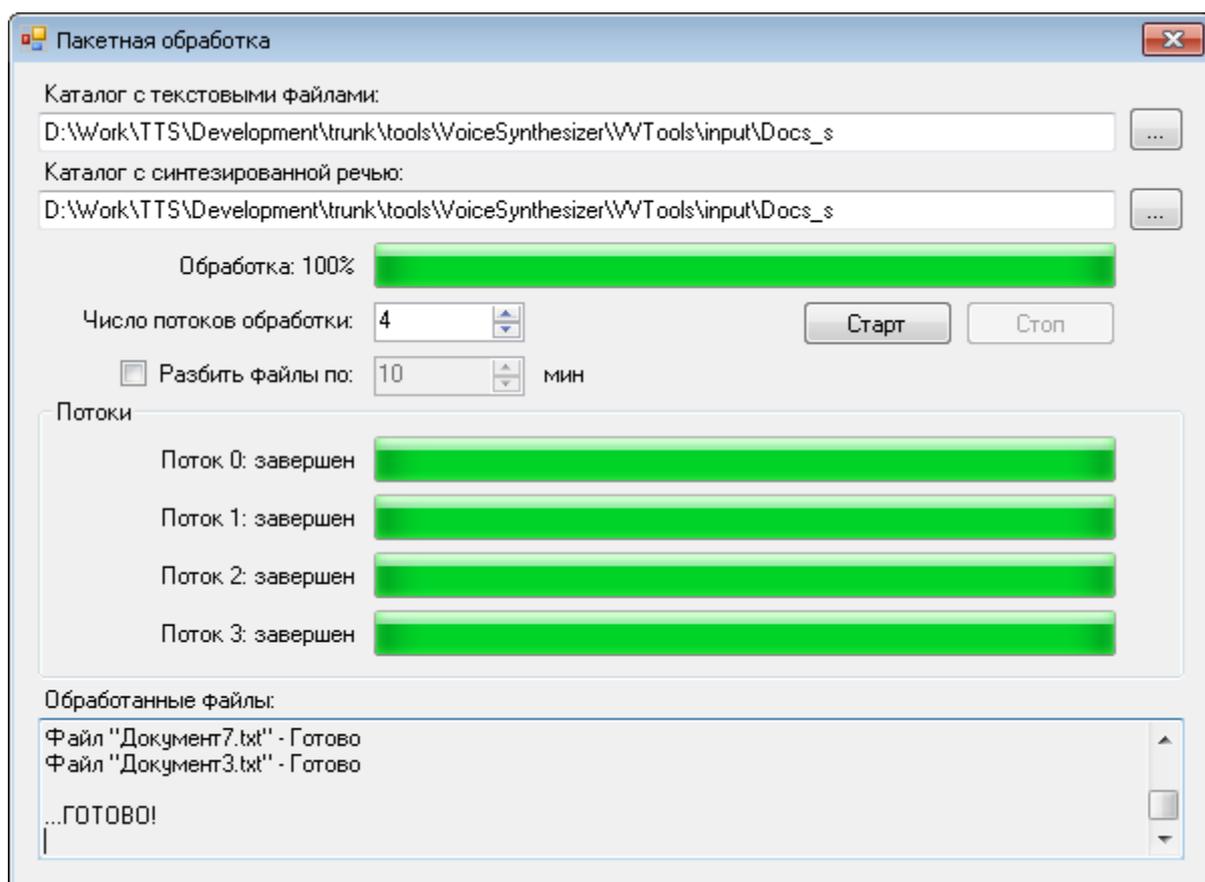


Рис. 2.1.9.2. Завершение процесса пакетной обработки.

Пользователь может прервать пакетную обработку, нажав кнопку «Стоп». При этом в результирующей папке будут сохранены те файлы, которые успели синтезироваться хотя бы частично.

Процесс пакетной обработки не запрашивает подтверждения на перезапись уже существующих файлов, забота о наличии файлов в каталоге с синтезированной речью возлагается на пользователя.

3. ПОЛЬЗОВАТЕЛЬСКИЙ СЛОВАРЬ

3.1 Функции пользовательского словаря

Редактор пользовательского словаря используется для решения следующих задач:

- Добавление в пользовательский словарь слов с указанием ударения, необходимого пользователю;
- Добавление в пользовательский словарь замен для различных последовательностей слов или символов (например, задание прочтения английских слов; расшифровка сокращений и аббревиатур; чтение тех или иных специальных символов);
- Установка знака пользовательского ударения, который затем может быть использован в тексте, подаваемом на синтез.

3.2 Общий интерфейс пользовательского словаря.

Основное окно пользовательского словаря имеет две вкладки. На первой вкладке «Словарь» (загружаемой по умолчанию) показано текущее содержание словаря: словарные статьи и замены (рис.3.1).

Для слов используются первый и третий столбцы таблицы. В первом столбце показана основа, в третьем столбце - через запятую все словоформы данной словарной статьи. Знаком «<» обозначено основное ударение, которое может находиться только после гласной. Например:

Хоп		хопѐ<р, хопра<, хопру<, хопро<м, хопре<
-----	--	---

Для замен используются первый и второй столбцы. В первом столбце – исходный текст, в последнем – его расшифровка. Например:

invade	инвэ<йд	
--------	---------	--

В левом нижнем углу расположено окно поиска слов. При пустом окне поиск производится по всему содержимому.

Далее, расположены переключатели регулирующие формат показа слов и замен. Оба переключателя могут быть одновременно выключены или включены поочередно, но одновременно не могут находиться в положении «включено».

Предусмотрена сортировка содержимого таблицы по любому из трех столбцов. По умолчанию, при загрузке производится сортировка по основам.

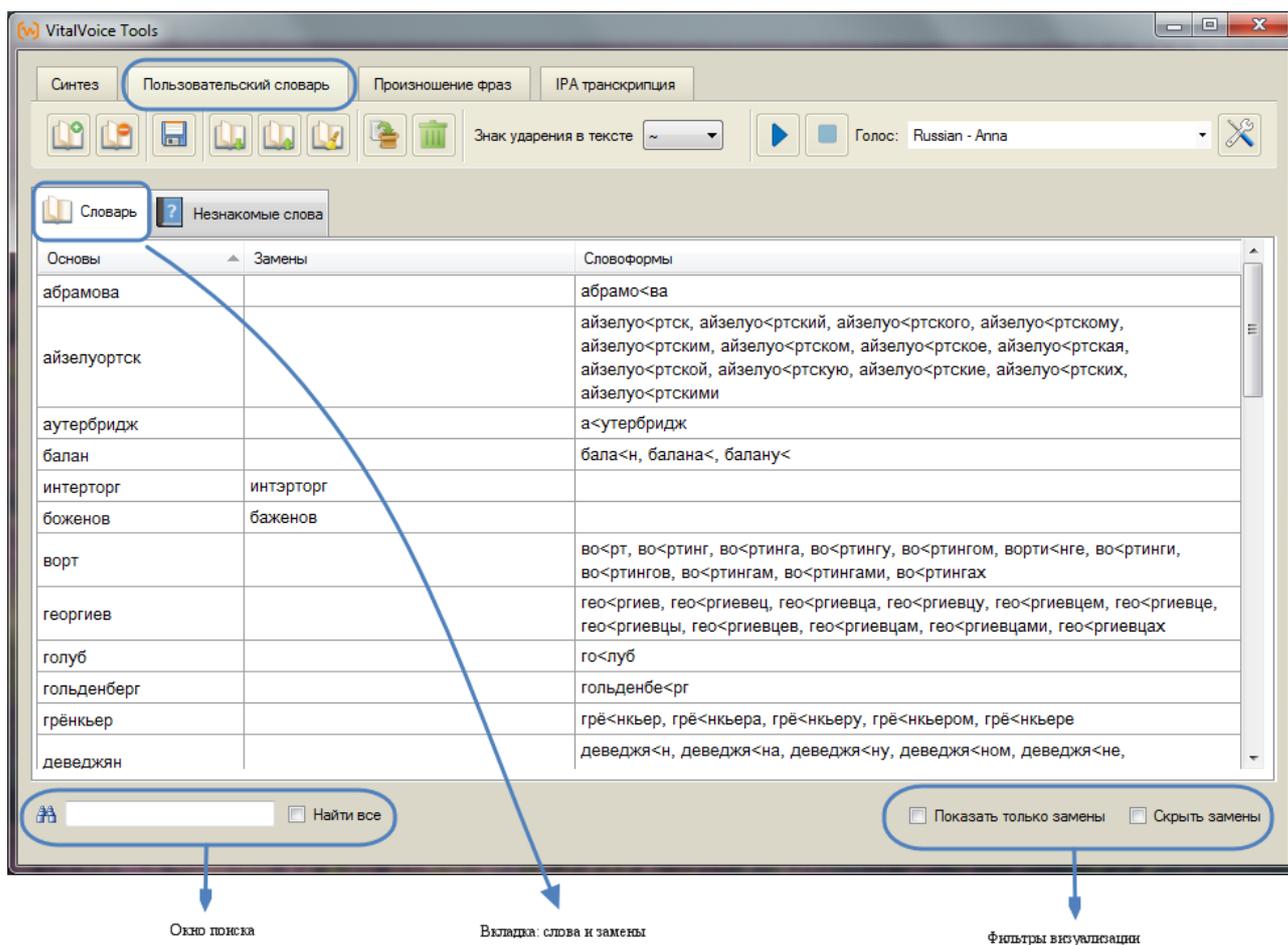


Рис.3.1. Вкладка «Словарь» основного окна пользовательского словаря

Удалить слова и замены можно выделив нужные элементы в таблице и нажав кнопку «Delete». Удаление элементов происходит по запросу (рис.3.2):

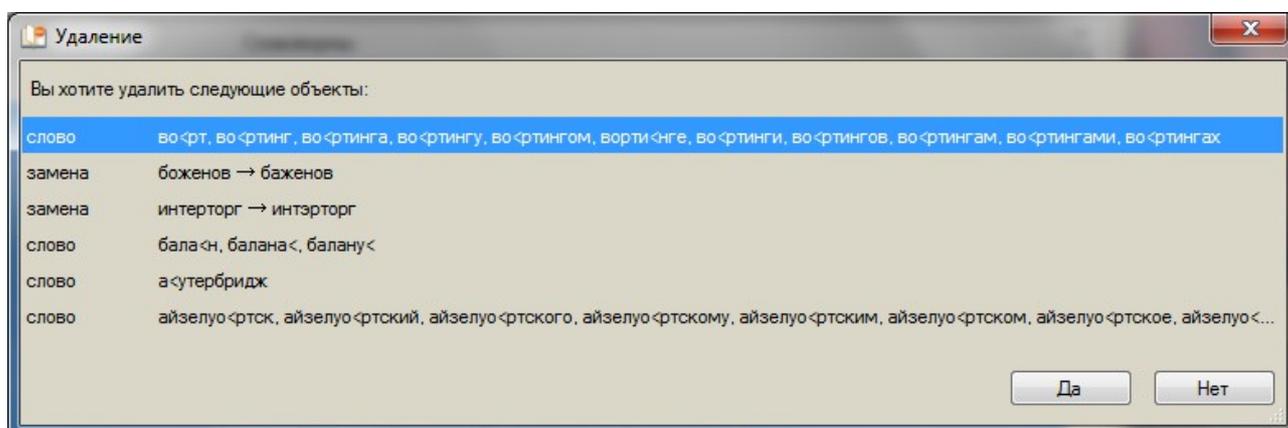


Рис.3.2.

На второй вкладке основного окна пользовательского словаря (рис.3.3) показаны русские и английские слова, которые не были найдены в словарях системы синтеза речи. Все слова расположены в алфавитном порядке. Также на этой вкладке можно использовать окно поиска.

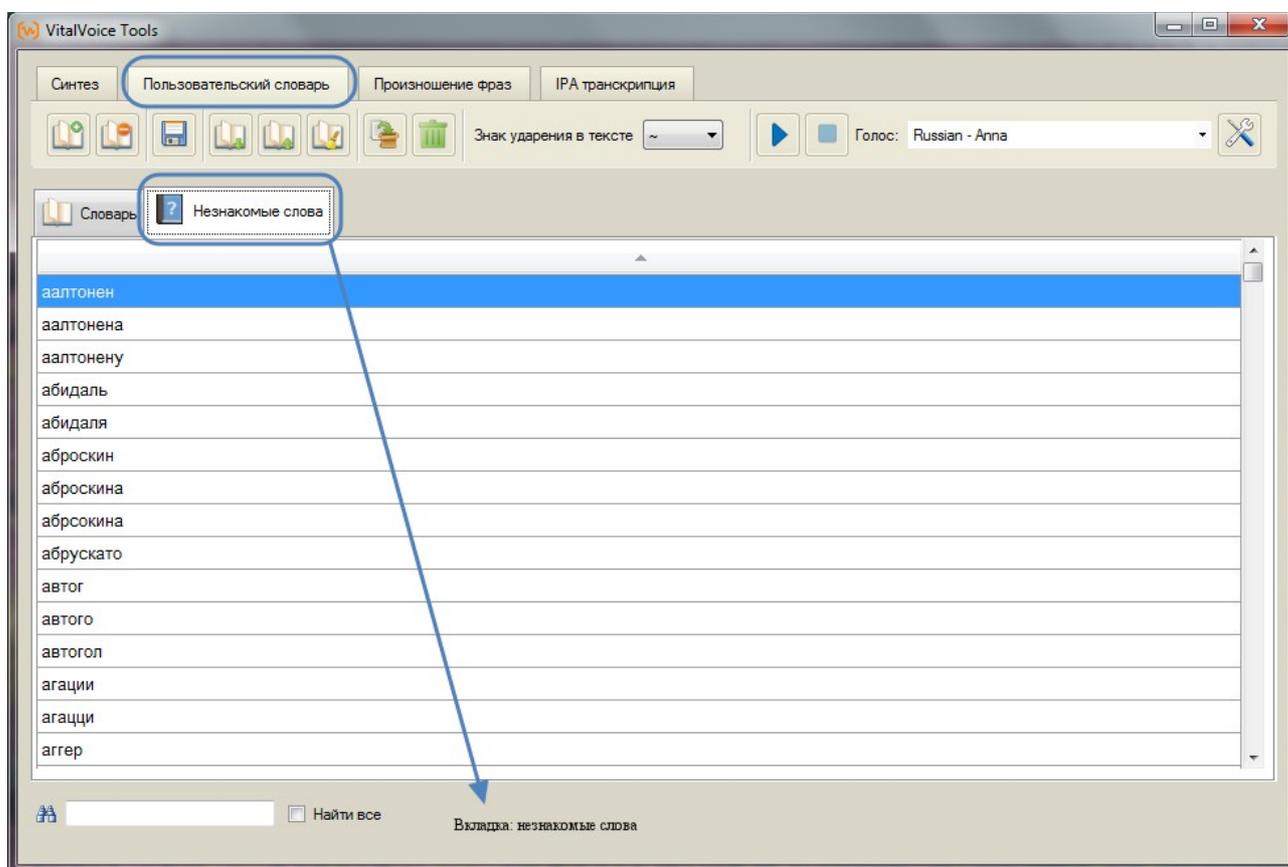


Рис.3.3. Вкладка «Незнакомые слова» окна пользовательского словаря

3.3 Добавление и редактирование слов и замен

Для добавления и редактирования слов и замен нужно нажать кнопку «Добавить слово» в верхней левой части главного окна (рис.3.4):

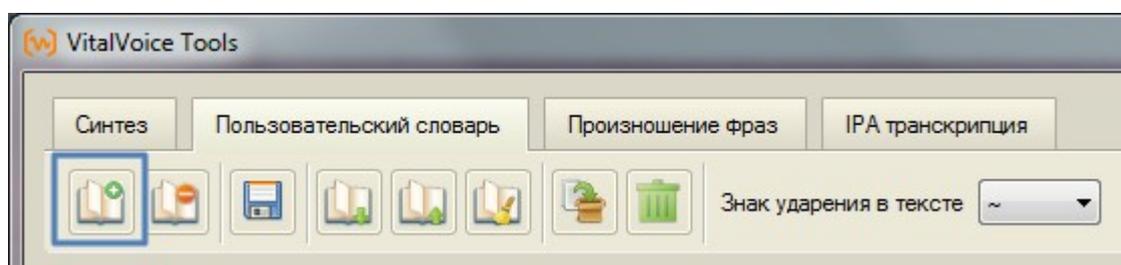


Рис.3.4.

Загрузится окно «Добавления и редактирования слов». Окно состоит из четырех вкладок (каждая для своего режима добавления и редактирования).

3.3.1 Добавление нового слова

1. Введите новое слово в поле редактирования на первой вкладке «**Новое слово**» окна «**Добавление и редактирование слова**» (рис.3.5). Не забудьте знак основного ударения (символ <)!

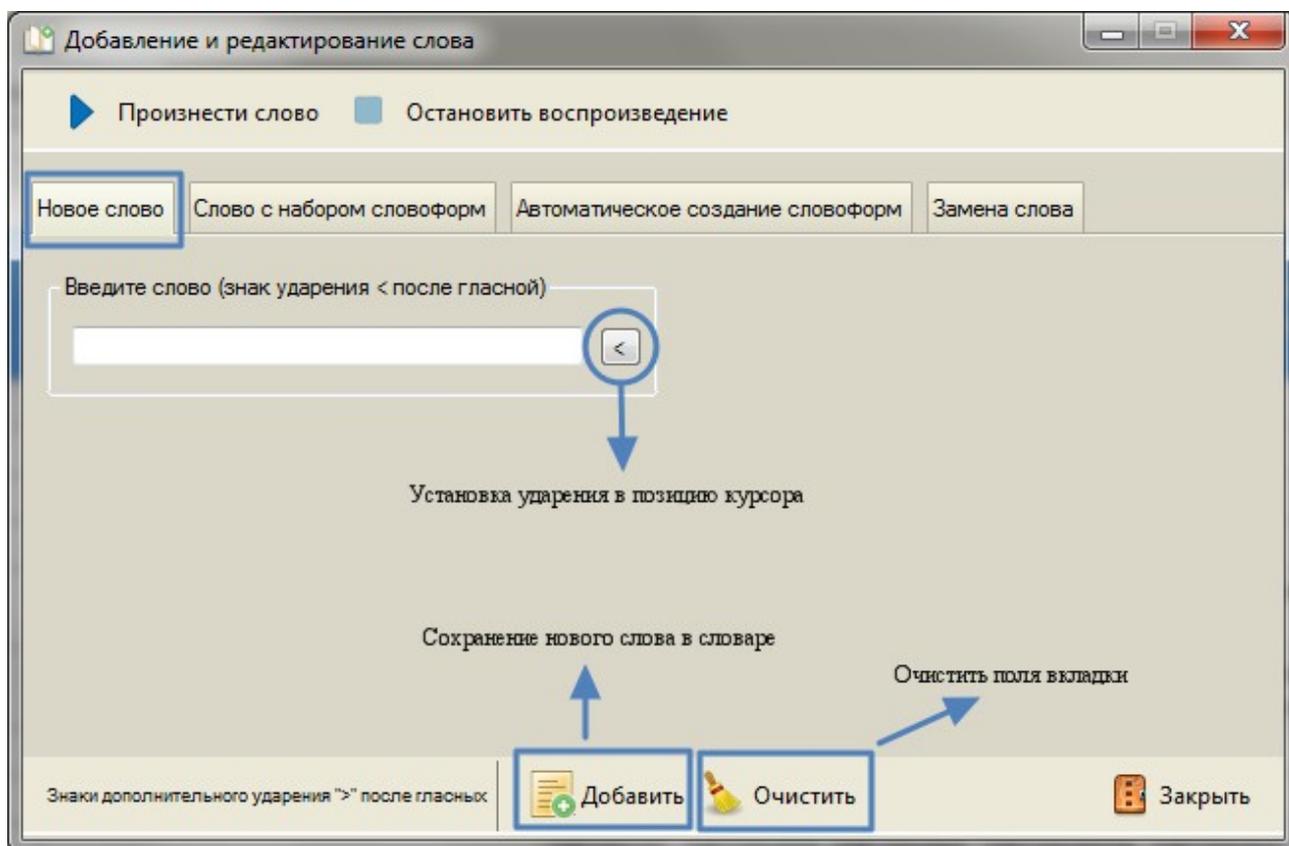


Рис.3.5.

2. Чтобы не переключать раскладку при установке знака ударения с русской на английскую, предусмотрена кнопка установки ударения в текущей позиции курсора.
3. Нажмите кнопку «**Сохранить**», находящуюся в нижнем левом углу окна. Слово будет отображено и выделено (голубым «LightSkyBlue») цветом в списке слов пользовательского словаря.
4. Если нужно ввести в словарь несколько слов, то повторите шаги 1-3 для добавления каждого слова.
5. Для выхода из окна без сохранения слова нажмите клавишу Esc.

3.3.2 Добавление слова с набором словоформ

Если вы хотите добавить словарную статью, включающую несколько форм одного и того же слова, то перейдите на закладку «Слово с набором форм» (рис.3.6).

1. Введите основу слова (ту часть слова, которая остается неизменной во всех формах) в поле «Основа». Например, если вы хотите добавить формы «Ива́нов, -а, -ы» введите ива<нов.

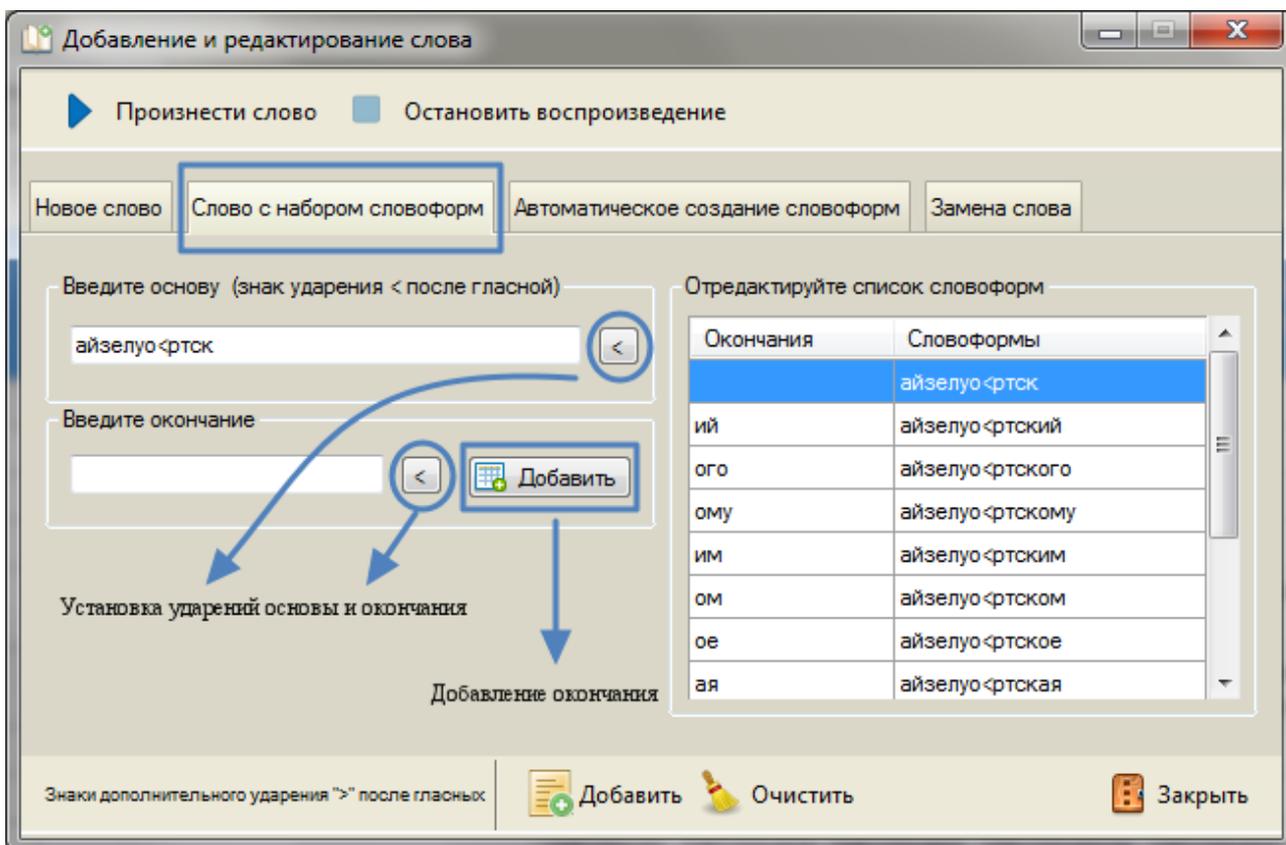


Рис.3.6.

2. Введите окончание словоформы в поле «Окончание». Например, для формы «Ива́нова» введите а.
3. Нажмите кнопку «Добавить». Словоформа появится в списке форм слова.
4. Для добавления формы с нулевым окончанием, нажмите кнопку «Добавить» при пустом поле «Окончание».
5. Не забудьте знак ударения (символ <)! Ударение может стоять либо на основе, либо на окончании. Ударение в окончании можно изменить, щелкнув левой кнопкой мышки по соответствующей строке окончания (рис.3.7).

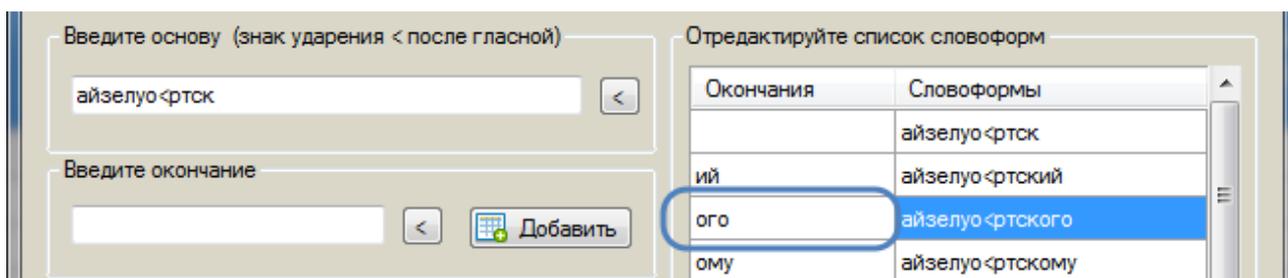


Рис.3.7.

Повторите шаги 2-5 для каждой из форм, которую вы хотите добавить в словарную статью.

Вы можете удалить из словарной статьи лишние формы, нажав клавишу «Delete». Будет выдан соответствующий запрос на удаление. Для добавления словарной статьи в словарь нажмите кнопку «Сохранить».

Для выхода из окна без сохранения слова нажмите клавишу Esc.

3.3.3 Автоматическое создание словоформ

Вкладка «Автоматическое создание словоформ» позволяет пользователю сохранить полную парадигму добавляемого в словарь слова, выбрав набор окончаний по образцу одной из парадигм, хранящихся в памяти словаря (рис.3.8).

Для добавления слова нужно выполнить следующие шаги:

1. Вести слово в начальной форме. Для существительного это именительный падеж единственного числа (например, "стол"), для прилагательного - именительный падеж единственного числа мужского рода (например, "красивый"), для глагола - неопределенная форма (инфинитив) (например, "делать"). Обязательно поставьте знак ударения. Для обозначения ударного гласного используется символ < (знак «меньше»), который ставится после ударной гласной. Вы также можете задать дополнительное ударение (одно или несколько) при помощи символа > (знак «больше»).

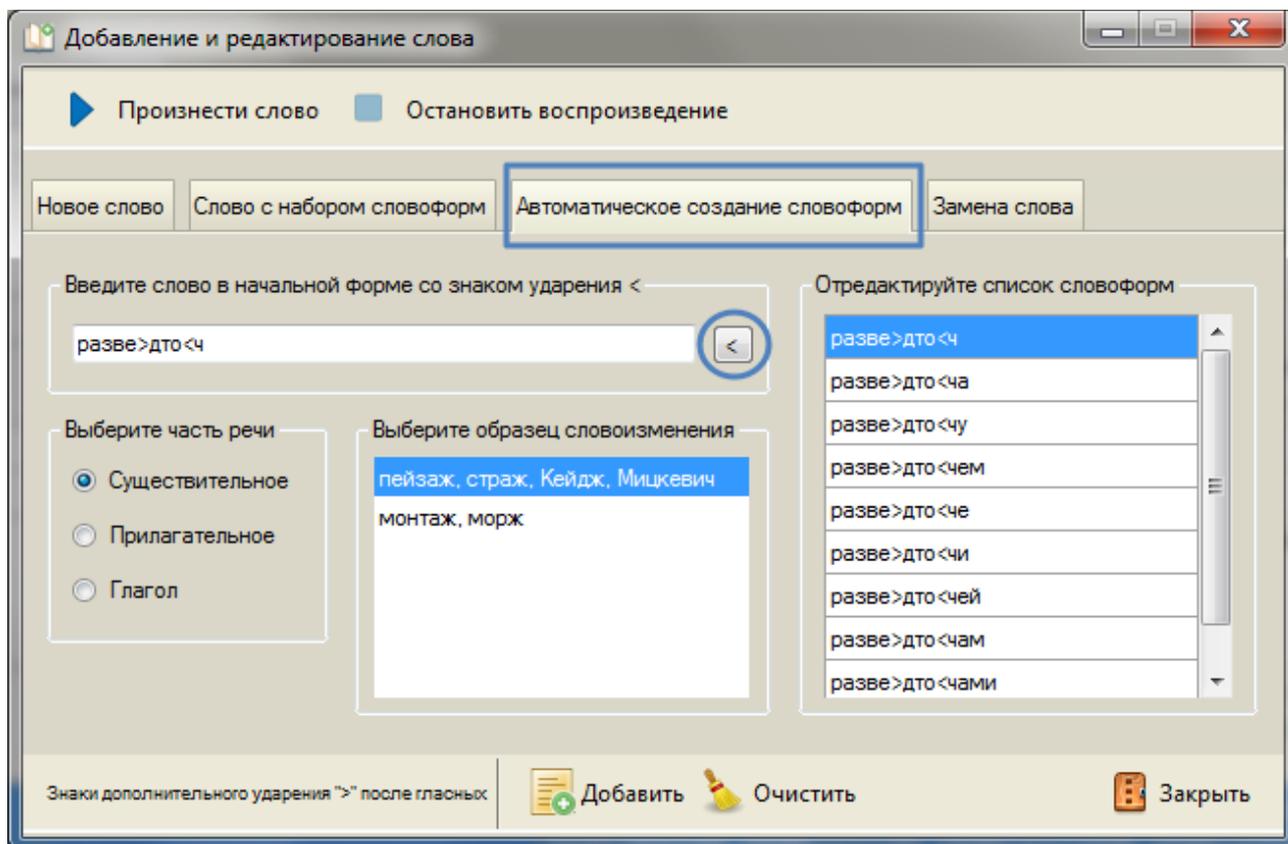


Рис.3.8.

2. Выберите часть речи вводимого слова (существительное, прилагательное или глагол). Если вы добавляете причастие, выберите прилагательное. Парадигма глагола не включает причастия, поэтому при необходимости добавить к глаголу причастия их нужно добавлять в словарь отдельно. Имена собственные, такие как фамилии, личные имена, географические названия, являются существительными.
3. Выберите образец словоизменения из тех, что представлены в списке. При однократном щелчке мышью по выбранному образцу, в таблице словоформ в правой части окна будет показано словоизменение вводимого слова по выбранной модели.
4. Проверьте правильность выбранной парадигмы. По умолчанию ударение во всех формах слова ставится на том же слоге, что и в начальной форме. Щелкнув два раза по форме в списке, вы можете изменить место ударения в конкретной форме (рис.3.9).

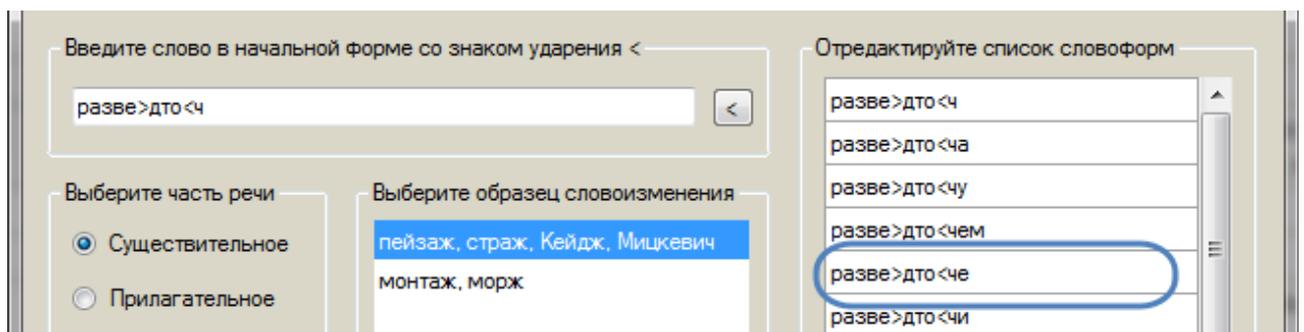


Рис.3.9.

5. Вы также можете удалить из парадигмы лишние формы, нажав клавишу «Delete». Будет выдан соответствующий запрос на удаление.
6. Нажмите кнопку "Сохранить" для завершения редактирования словарной статьи. Слово будет добавлено в пользовательский словарь.

3.3.4 Замена слова

Вкладка «Замена слова» имеет следующий вид (рис.3.10):

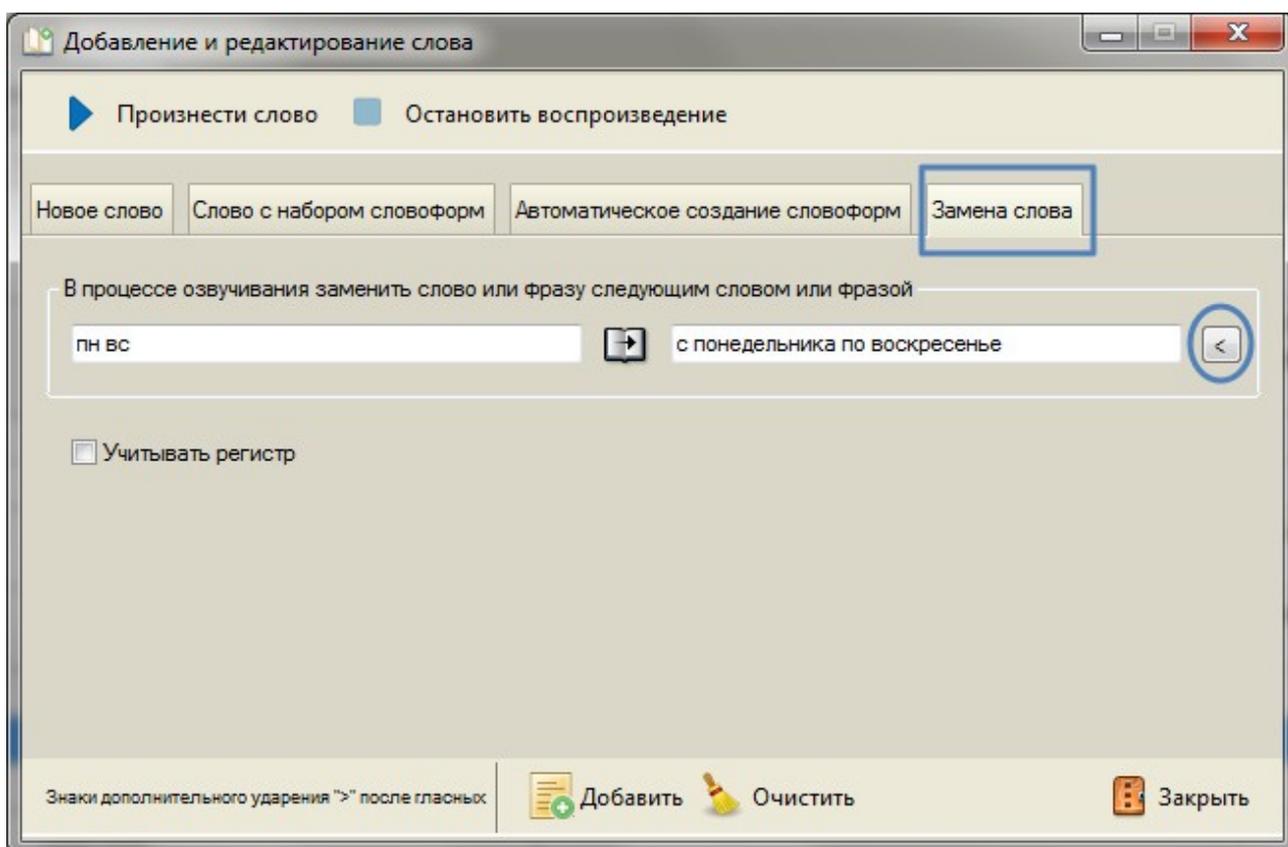


Рис.3.10.

С помощью этой вкладки вы можете:

- задать расшифровку сокращений или аббревиатур;
- задать чтение слов, написанных латиницей;
- задать чтение специальных символов

Последовательность действий следующая:

1. Введите текст в поле для исходного текста. Вы можете использовать также знаки препинания и другие символы. Например, если нужно расшифровать «пн. - вс.» как «с понедельника по воскресенье», то введите **пн. - вс.**
2. Введите желаемую расшифровку в поле замены. Например, **с понедельника по воскресенье.**
3. Вы можете задать желаемые ударения в расшифровке, используя символы <.
4. Нажмите кнопку «**Сохранить**».
5. Если нужно ввести в словарь несколько замен, то повторите шаги 1-3 для добавления каждого из них.

3.3.5 Редактирование слов и замен

Для редактирования содержащихся в словаре слов или замен дважды щелкните мышью на соответствующей строке в списке главного окна.

Для элемента **слово** окно редактирования откроется на второй вкладке «**Слово с набором словоформ**». На этой вкладке уже будет загружено слово со всеми словоформами. Кроме того, это слово будет загружено и готово к редактированию в первой (как одиночное слова) и в третьей (автоматическое создание словоформ) вкладках.

Для элемента **замена** окно редактирования откроется на последней вкладке «**Замена слова**». В текстовые поля будет загружены исходный текст и текст замены.

3.3.6 Ударение в слове

Ударение в слове задается символом < (знак «меньше»), который ставится после ударной гласной. Если в слове не задано ударение, слово в словаре сохранено не будет, будет выдано предупреждение «В слове нет знака ударения».

Вы также можете задать в слове дополнительные ударения при помощи символа > (знак «больше»).

В слове может быть только одно основное ударение, но дополнительных ударений может быть несколько. При наличии дополнительных ударений в слове также обязательно должно быть основное ударение. При попытке сохранить слово, в котором есть дополнительное ударение (одно или несколько), но нет основного ударения, будет выдано предупреждение «В слове нет основного ударения».

Ударение может быть поставлено только после гласной буквы. При попытке поставить знак ударения после согласной знак печататься не будет.

3.4 Знак ударения в тексте

Вы можете выбрать знак ударения из списка в поле «**Знак ударения в тексте**» (рис.3.11).

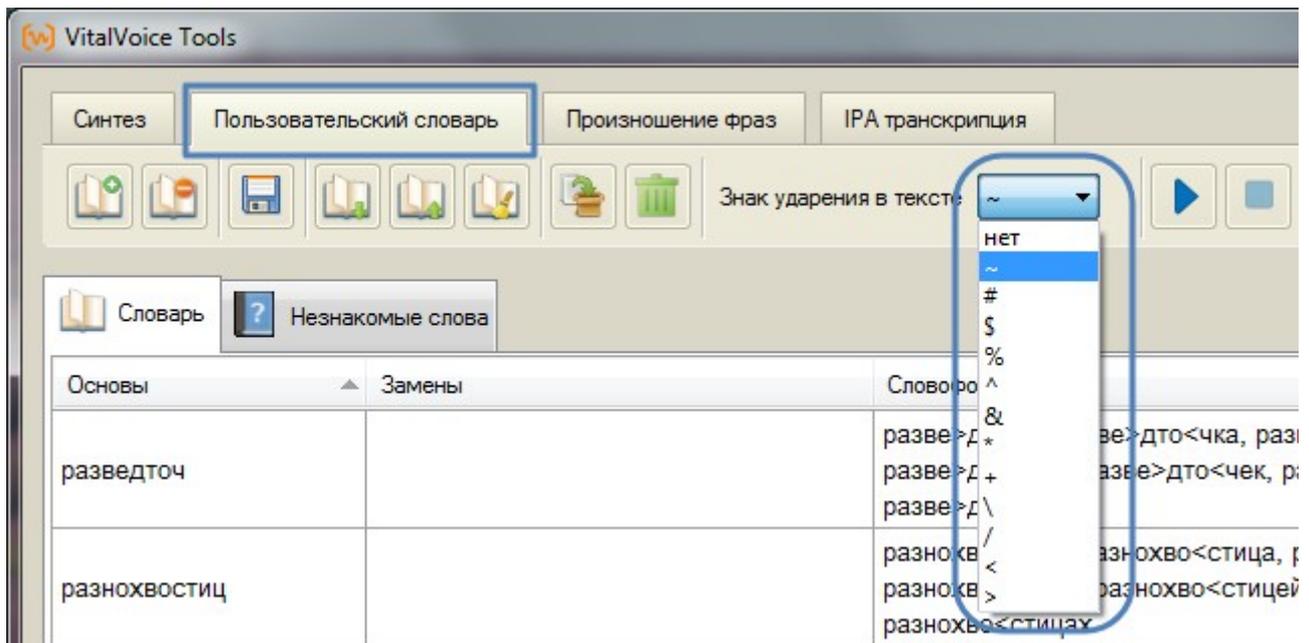


Рис.3.11.

Если один из предоставленных знаков выбран в качестве знака пользовательского ударения, то этот знак может быть использован для изменения ударения в слове при подаче текста на синтез речи.

Например, если выбран знак ~ то при чтении текста:

"по до~роге едет тракто~р." синтезатор произнесет: "по дОроге едет трактОр."

Если в поле «Знак ударения» выбрано значение «Нет», то не существует знака, который мог бы быть использован для изменения ударения в слове при подаче текста на синтез речи.

3.5 Инструменты

Панель инструментов находится в верхней части главного окна пользовательского словаря (рис.3.12).

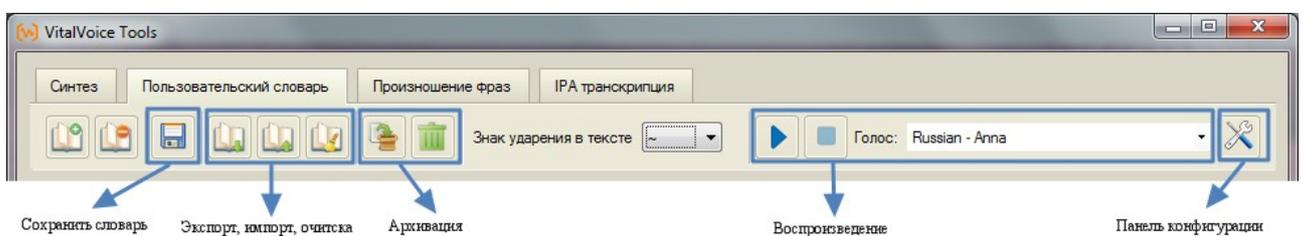


Рис.3.12.

- **Экспорт словаря** позволяет сохранить текущий словарь слов и замен как текстовый файл по заданному пользователем маршруту.

- При **импорте словаря** словарь по заданному пользователем маршруту будет загружен в главное окно.
- При **очистке словаря** список слов в главном окне будет очищен.
- Кнопка «**архивация**» позволяет сохранить текущее состояние словаря слов и замен по стандартному для движка синтеза маршруту в поддиректорию **UserDictArchive** (см. раздел «**Загрузка словарей в начале работы программы**»).

Окно конфигурации имеет следующий вид (рис.3.13):

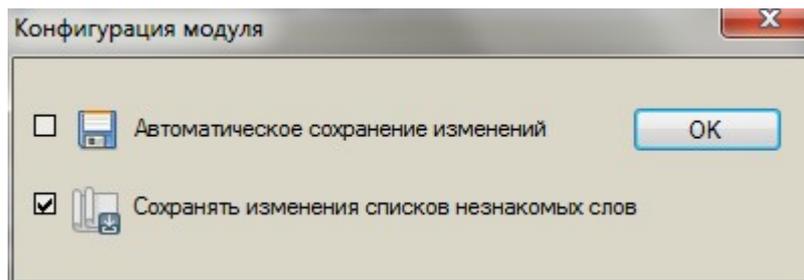


Рис.3.13.

3.6 Загрузка словарей в начале работы программы

Файл пользовательского словаря (**Dict.UserWords.dic**) и файлы незнакомых русских (**UnknownWordList.txt**) и английских (**UnknownEngWordList.txt**) слов хранятся по маршруту, заданному в реестре:

HKEY_LOCAL_MACHINE\SOFTWARE\Speech Technology Center\TTS\Languages\Russian

Файл пользовательского словаря представляет собой текстовый файл и доступен для редактирования вручную. Инструкция по редактированию файла содержится в теле файла.

Файлы незнакомых русских и английских слов модифицируются движком синтеза в режиме дозаписи, поэтому без коррекции могут достигать больших объемов. В такой ситуации в начале загрузки будет выдано предупреждение (рис.3.14):

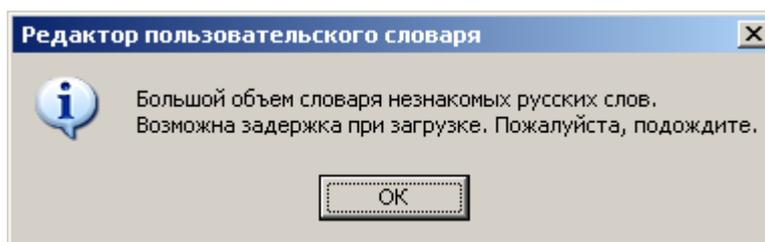


Рис.3.14.

При этом соответствующий словарь незнакомых слов будет сформирован из последних 1000 слов.

3.7 Данные пользовательского словаря в системе синтеза речи

3.7.1 Чтение слова с учетом данных пользовательского словаря

При наличии слова в пользовательском словаре, ударение в данном слове всегда будет проставлено так, как оно задано в пользовательском словаре.

При наличии в написании слова в пользовательском словаре буквы Ё, данное слово будет прочитано с Ё, даже если в тексте оно будет написано через Е.

Обратите внимание, что слово, добавленное в пользовательский словарь, будет читаться с желаемым ударением только в тех формах, в которых оно сохранено в словаре. Например, если в пользовательский словарь добавлено слово «ива<нов» (с ударением на «а»), и это единственная сохраненная в пользовательском словаре форма, то форма «иванова» будет прочитана по умолчанию (с ударением на «о»). Чтобы все формы слова читались с желаемым ударением, нужно все их добавить в словарь.

После сохранения изменений в пользовательском словаре необходимо перезагрузить программу синтеза речи, чтобы изменения вступили в силу!

3.7.2 Обработка пользовательских замен

Обработка пользовательских замен и расшифровок при синтезе речи производится с учетом заданных знаков препинания. Например, если задана расшифровка «пн. - вс.» как «с понедельника по воскресенье», то написание «пн-вс» не будет расшифровано заданным образом. Этот вариант следует ввести отдельно.

Знаки препинания, стоящие в поданном на синтез тексте перед или после заданного для расшифровки элемента, будут обработаны в обычном режиме.

Если заданный для расшифровки элемент будет разделен в тексте знаком конца абзаца (переводом каретки), то расшифровки не произойдет. Так, в рассмотренном примере следующий текст не будет расшифрован заданным образом:

пн.-
пт.

Обработка пользовательских замен производится без учета регистра. Так, если в словарь введена замена слова «anne» на «а<нна», то в тексте слово «Anne» будет также прочитано.

3.7.3 Обработка знака пользовательского ударения в тексте

Знак пользовательского ударения, заданный через словарь пользователя, может быть затем использован в тексте, подаваемом на синтез, для задания ударения в отдельных словах.

Знак пользовательского ударения должен стоять в тексте после гласной. Например, если задан знак пользовательского ударения ~, то при подаче на синтез текста «Здравствуйте, господин Ива~нов» слово «Иванов» будет прочитано с ударением на «а».

Если знак пользовательского ударения не определен, то при подаче текста на синтез нельзя задать ударение с помощью знака после гласной. В этом случае для изменения ударения в синтезируемом тексте можно использовать соответствующий тег.

4. ПРОИЗНОШЕНИЕ ФРАЗ

4.1. Назначение

Модуль «Произношение фраз» (ПФ) применяется для настройки пользователем произношения отдельных фраз путем модификации транскрипции (вставка / подмена / удаление монофонов) и управления процедурой подбора звуковых единиц методом Unit Selection (запрет нежелательных кандидатов / фиксация выбранного кандидата).

4.2. Хранение данных

Произношения фраз хранятся в базе данных пользовательских произношений (далее – БД ПП). Произношение каждой отдельной фразы может быть настроено только для конкретного диктора (точнее – для конкретной голосовой базы): единицей хранения данных является пара «фраза + голосовая база». Голосовая база идентифицируется по признаку: путь к файлу voice.dat. При необходимости настройки одной и той же фразы для нескольких дикторов (для разных голосовых баз) эту фразу потребуется настраивать для каждой базы отдельно.¹ Это связано с особенностью организации голосовой базы – в каждой базе имеется свой набор звуковых единиц, которые отличаются по своим параметрам, и могут быть не одинаковы для разных баз, и обязательно различаются для разных дикторов.

БД ПП хранится в файле Transform.Phrase.Database.txt, который находится в папке с пользовательскими данными, путь к которой определяется ключом реестра

```
HKLM\SOFTWARE\Speech Technology Center\TTS\Languages\<...>\UserDataPath
```

где <...> = Russian для русского языка. При отсутствии данного ключа реестра БД ПП будет сохраняться в текущей директории.

4.3. Использование совместно с движком синтеза (TTS Engine)

Для подключения БД ПП к движку синтеза требуется задать ключ конфигурации

```
("TTS/phonetic/transform/user_phrases", "on")
```

При этом для каждой фразы, найденной в БД ПП для используемого Синтезом голоса, вместо стандартного произношения будет подставлено созданное пользователем произношение, хранящееся в БД ПП.

¹ Для голосов одного и того же диктора, имеющих совпадающие наборы звуковых единиц, но разную частоту дискретизации, на уровне движка TTS реализована возможность поиска фразы того же диктора с другой частотой дискретизации. Для обеспечения возможности такого поиска пути, по которым установлены голосовые базы, должны совпадать с точностью до указанной частоты, причем она должна быть указана в конце строки, например [c:\path\Anna8000] и [c:\path\Anna].

4.4. Интерфейс и использование модуля

4.4.1. Условия работы модуля

Для работы модуля требуется наличие установленного движка синтеза (TTS Engine) и голосов, совместимых с этим движком. Редактирование произношения фраз доступно только для установленных в системе голосов.

4.4.2. Запуск модуля

Модуль запускается при первом переходе на соответствующую вкладку основного окна приложения. При запуске происходит загрузка БД ПП из папки с пользовательскими данными (см. разд. 4.2. *Хранение данных*).

4.4.3. Интерфейс модуля

Внешний вид интерфейса модуля изображен на *рис. 4.4.3*.

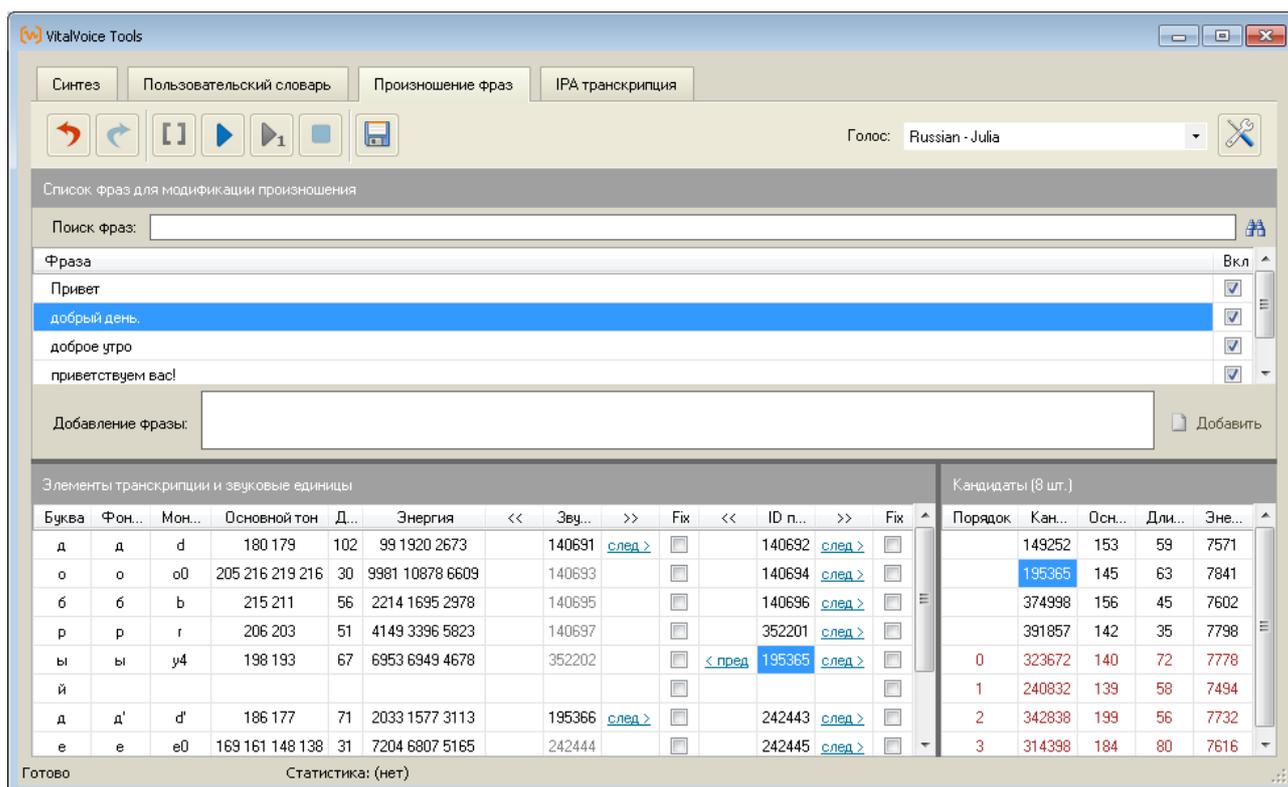


Рис. 4.4.3. Внешний вид интерфейса модуля «Произношение фраз».

Кроме того, одним из основных элементов интерфейса приложения является диалоговое окно выбора монофона (см. разд. 4.5.4. *Модификация транскрипции*), а также диалоговое окно конфигурации модуля (см. разд. 4.4.3.2. *Диалог конфигурации модуля*)

Подробнее о редактировании транскрипции см. разд. 4.5.4. Модификация транскрипции.

4.4.3.1. Панель инструментов

Панель инструментов расположена снизу от главного меню. Кнопки на панели инструментов позволяют производить следующие операции:

Отменить сделанную операцию (Undo). Отменяет очередное изменение произношения выбранной фразы, откатываясь на один шаг назад по истории изменений. *Подробнее об истории изменений см. разд. 4.5.6. История изменений (Undo).*

Вернуть сделанную операцию (Redo). Возвращает очередное отмененное изменение произношения выбранной фразы, продвигаясь на один шаг вперед по истории изменений. *Подробнее об истории изменений см. разд. 4.5.6. История изменений (Undo).*

Генерация транскрипции (□). Позволяет сгенерировать транскрипцию по умолчанию для выбранной фразы. *Подробнее о редактировании транскрипции см. разд. 4.5.4. Модификация транскрипции.*

Произнести фразу (Play). Произносит фразу с применением настроенного пользовательского произношения. *Подробнее о произнесении фраз см. разд. 4.5.7. Произнесение фраз (Play).*

Произношение по умолчанию (Play1). Произносит фразу с применением произношения по умолчанию (без использования БД ПП). *Подробнее о произнесении фраз см. разд. 4.5.7. Произнесение фраз (Play).*

Стоп (Stop). Прервать произнесение фразы.

Сохранить (Save). Сохраняет БД ПП в файл. *Имя и расположение файла описаны в разд. 4.2. Хранение данных.*

Голос и выпадающий список справа от него позволяют выбрать один из голосов, установленных в системе, для которого будут редактироваться произношения фраз. При смене голоса отображается список фраз, хранящихся в БД ПП для данного голоса.

Настройка (Tools). Показывает диалоговое окно конфигурации модуля (*рис. 4.4.3.1*).

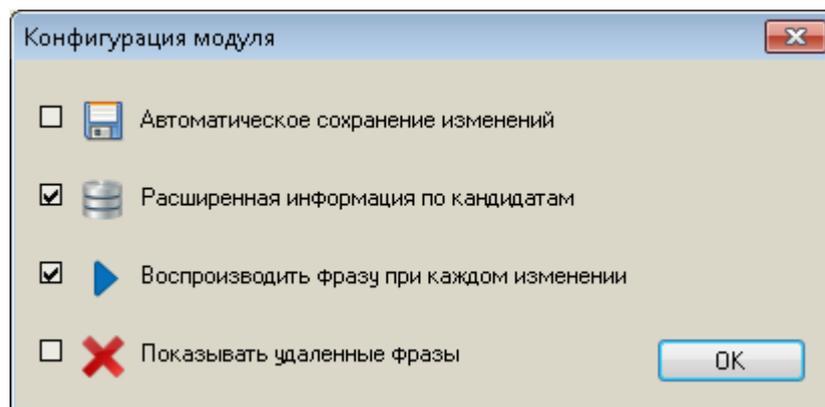


Рис. 4.4.3.1. Диалог конфигурации модуля.

4.4.3.2. Диалог конфигурации модуля

Диалоговое окно конфигурации предназначено для изменения параметров конфигурации модуля. Возможно изменение следующих параметров.

Автоматическое сохранение изменений (Save). Галочка. Включает / отключает автоматическое сохранение в файл каждого изменения, внесенного пользователем в произношение фраз.

Расширенная информация по кандидатам (DB). Галочка. Разрешает / запрещает загрузку из голосовой БД информации по кандидатам. *Подробнее – см. разд. 4.4.3.5. Таблица кандидатов.*

Произносить фразу при каждом изменении (Play). Галочка. Включает / отключает произношение фразы после каждой операции редактирования (вставка / удаление монофона, запрет кандидата, ...) *Подробнее – см. разд. 4.5.4. Модификация транскрипции, 4.5.5. Подбор звуковых единиц.*

Показывать удаленные фразы (Delete). Галочка. Включает / отключает отображение удаленных фраз. Удаленные фразы показываются как неиспользуемые. *Подробнее – см. разд. 4.5.2. Создание и редактирование фразы.*

4.4.3.3. Таблица фраз

Табличное представление «**Список фраз для модификации произношения**» в верхней части окна под панелью инструментов содержит список фраз, имеющих (созданных) для выбранного голоса.

Ячейки столбца «**Фраза**» содержат текстовое представление фразы. Редактирование фразы производится прямо в ячейке.

Ячейки столбца «**Вкл**» позволяют включать / отключать использование фразы в движке синтеза.

Подробнее о работе с фразами см. разд. 4.5.2. Создание и редактирование фразы.

4.4.3.4. Таблица элементов транскрипции

Табличное представление «**Элементы транскрипции и звуковые единицы**» в левой нижней части окна содержит список элементов транскрипции и подобранных звуковых единиц для выбранной фразы.

Ячейки столбца «**Буква**» содержат буквы, из которых состоит фраза. Паузы обозначаются символом «#».

Ячейки столбца «**Фонема**» содержат фонемы, представляющие транскрипцию фразы, записанную кириллицей.

Ячейки столбца «**Монофон**» содержат монофоны, представляющие фонемную транскрипцию фразы, записанную с помощью принятых в TTS обозначений.

Ячейки столбцов «**Основной тон**», «**Длительность**» и «**Энергия**» содержат значения соответствующих параметров, предсказанные для соответствующих звуковых элементов фразы.

Ячейки столбцов «**Звуковые единицы (ID)**» и «**ID правого полуфона**» содержат идентификаторы звуковых единиц, выбранных методом Unit Selection для озвучивания соответствующих монофонов.

Ячейки столбцов «**Fix**» позволяют зафиксировать выбранные звуковые единицы для озвучивания соответствующего монофона.

Заголовки столбцов «Fix». Нажатие на заголовок столбца позволяет зафиксировать произношение всех звуковых единиц выбранной фразы. После фиксации произношения выбранная фраза для выбранного голоса всегда будет звучать так, как указано пользователем, даже при возможном изменении алгоритма подбора звуковых элементов.

Подробнее о работе с элементами транскрипции и настройке процесса подбора звуковых единиц см. разд. 4.5.4. Модификация транскрипции, 4.5.5. Подбор звуковых единиц.

Дополнительно об алгоритме подбора звуковых элементов см. разд. 4.6.2. Особенности подбора звуковых элементов. Применительно к фиксации

произношения фразы – см. разд. 4.6.3. Изменения алгоритма подбора элементов.

4.4.3.5. Таблица кандидатов

Табличное представление «**Кандидаты**» в правой нижней части окна содержит список кандидатов, участвующих в подборе звуковых единиц методом Unit Selection для выбранного монофона.

Ячейки столбца «**Порядок**» для тех кандидатов, выбор которых запрещен для озвучивания соответствующих монофонов, содержат число, соответствующее порядку их запрещения при подборе кандидатов. Для тех кандидатов, выбор которых не был запрещен, ячейки этого столбца пусты.

Ячейки столбца «**Кандидаты (ID)**» содержат идентификаторы звуковых единиц – кандидатов, участвующих при подборе звуковых единиц методом Unit Selection.

Ячейки столбца «**Основной тон**» содержат значения основного тона соответствующего кандидата.

Ячейки столбца «**Длительность**» содержат значения длительности соответствующего кандидата.

Ячейки столбца «**Энергия**» содержат значения энергии соответствующего кандидата.

Подробнее о работе с кандидатами см. разд. 4.5.5. Подбор звуковых единиц.

4.5. Работа с приложением

4.5.1. Выбор голоса

Выбор голоса, для которого будут создаваться фразы и будет настраиваться их произношение, производится путем выбора требуемого элемента выпадающего списка в верхней части окна справа от пункта меню «Голос» (*см. разд. 4.4.3.1. Панель инструментов*).

При смене выбранного голоса в таблице фраз (*см. разд. 4.4.3.3. Таблица фраз*) отображаются фразы, соответствующие выбранному голосу.

4.5.2. Создание и редактирование фразы

Редактирование текстового представления фразы производится в ячейке таблицы фраз (см. разд. 4.4.3.3. *Таблица фраз*) в столбце «Фраза» путем нажатия клавиши F2, либо левого щелчка по текущей ячейке, либо двойного щелчка по ячейке. При изменении текста фразы производится автоматическая генерация транскрипции (см. разд. 4.5.4. *Модификация транскрипции*) с сохранением истории изменений (см. разд. 4.5.6. *История изменений (Undo)*). Несущественные изменения, такие как добавление знаков пунктуации, не вызывают регенерацию транскрипции, однако при желании ее можно сгенерировать принудительно, нажав кнопку «Генерация транскрипции» ([]) на панели инструментов (см. разд. 4.4.3.1. *Панель инструментов*).

Создание новой фразы производится путем ввода текста требуемой фразы в строку добавления новой фразы и нажатия кнопки «Добавить» справа от нее (см. разд. 4.4.3.3. *Таблица фраз*).

Удаление фразы производится путем выбора пункта «Удалить фразу» контекстного меню, вызываемого правым щелчком по ячейке с текстом фразы. В модуле ПФ реализовано логическое удаление фразы путем пометки ее как неиспользуемой, без физического удаления из БД ПП. Для восстановления удаленной фразы следует включить отображение удаленных фраз, включив галочку «Показывать удаленные фразы» (Check Delete) в окошке настройки конфигурации модуля (см. разд. 4.4.3.2. *Диалог конфигурации модуля*) и затем заново включить использование удаленной фразы. Физическое удаление фразы из БД ПП возможно путем удаления неиспользуемой фразы в режиме отображения удаленных фраз. Физическое удаление фразы не может быть отменено средствами, предоставляемыми модулем ПФ.

Галочка в ячейке столбца «Вкл» позволяет включать / выключать использование соответствующей фразы в движке, аналогично логическому удалению фразы.

4.5.3. Настройка произношения фразы

Модуль ПФ предоставляет возможность настройки произношения фразы путем модификации ее транскрипции, а также путем влияния на процесс подбора звуковых единиц методом Unit Selection.

4.5.4. Модификация транскрипции

Изначально при создании фразы для нее генерируется транскрипция по умолчанию. Транскрипцию фразы можно модифицировать, добавляя, удаляя или заменяя монофоны, а также добавляя и удаляя паузы.

Сгенерировать для фразы транскрипцию по умолчанию можно и принудительно, нажав кнопку «Генерация транскрипции» ([]) на панели инструментов (рис. 4.5.4.1) (см. разд. 4.4.3.1. Панель инструментов).

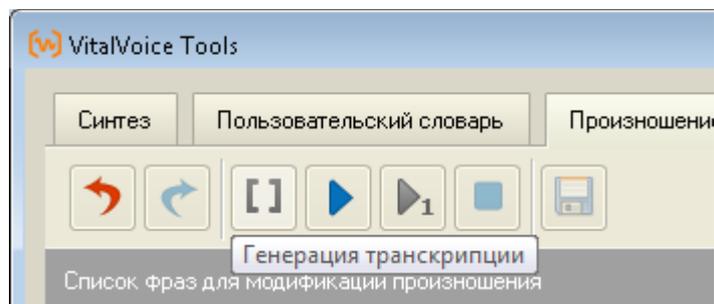


Рис. 4.5.4.1. Генерация транскрипции по умолчанию.

Добавление или вставка монофона производится путем выбора пунктов «Добавить монофон» или «Вставить монофон» контекстного меню, вызываемого правым щелчком по ячейке с буквой, фонемой или монофоном (рис. 4.5.4.2). «Добавить монофон» означает добавление монофона после текущей строки, «Вставить монофон» означает вставку монофона перед текущей строкой. После этого следует в появившемся окне (рис. 4.5.4.3) путем нажатия на соответствующую кнопку выбрать требуемый монофон и нажать ОК.

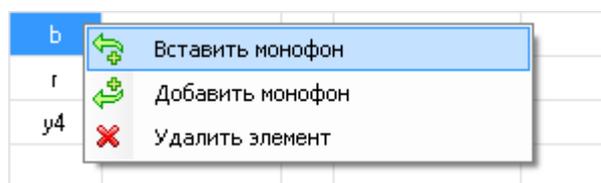


Рис. 4.5.4.2. Добавить монофон.

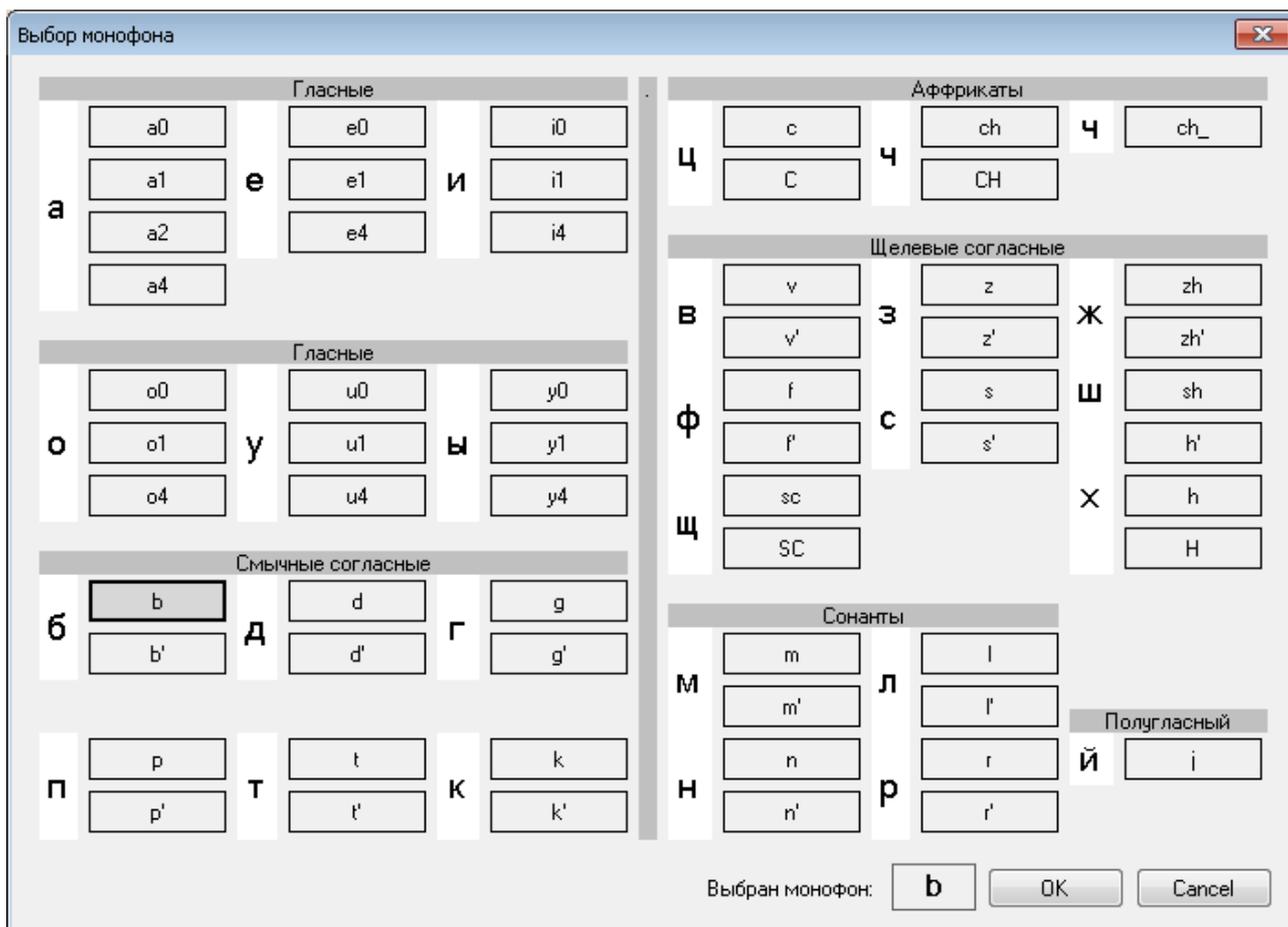


Рис. 4.5.4.3. Диалоговое окно выбора монофона.

Замена монофона производится путем двойного щелчка по ячейке с буквой, фонемой или монофоном, либо путем начала редактирования одной из этих ячеек (нажатие F2, либо щелчок по текущей ячейке, либо двойной щелчок по ячейке). После этого следует выбрать в появившемся окне (рис. 4.5.4.3) требуемый монофон и нажать ОК.

Добавление или вставка паузы производится путем выбора пунктов «Добавить паузу» или «Вставить паузу» контекстного меню, вызываемого правым щелчком по ячейке с буквой, фонемой или монофоном (рис. 4.5.4.4). «Добавить паузу» означает добавление паузы после текущей строки, «Вставить паузу» означает вставку паузы перед текущей строкой. Существует пять типов паузы, различающихся продолжительностью. Пауза может быть добавлена только на границе слов, и только в том случае, если пауза в этом месте отсутствовала.

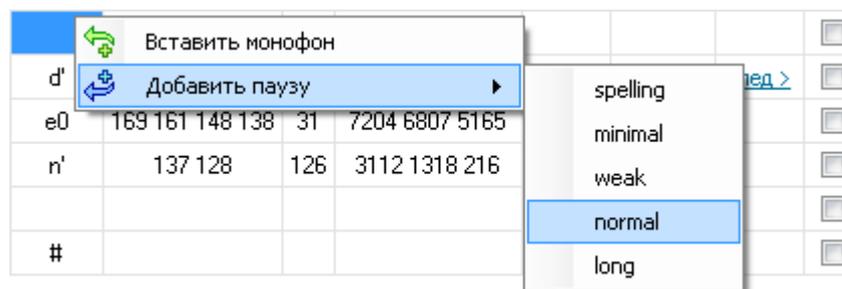


Рис. 4.5.4.4. Добавить паузу.

Удаление монофона или паузы производится путем выбора пункта «Удалить элемент» контекстного меню, вызываемого правым щелчком по ячейке с буквой, фонемой или монофоном.

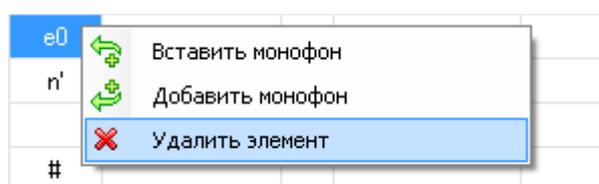


Рис. 4.5.4.5. Удалить элемент.

После любой из операций редактирования транскрипции производится либо только подбор звуковых единиц для озвучивания фразы, либо произнесение фразы – в зависимости от состояния галочки «Произносить фразу при каждом изменении» (см. разд. 4.4.3.2. Диалог конфигурации модуля).

4.5.5. Подбор звуковых единиц

Настройка процесса подбора звуковых единиц возможна путем запрета выбора отдельных кандидатов или фиксации выбранных кандидатов.

Запрет выбора кандидата для озвучивания монофона производится путем выбора пункта «Запретить данный кандидат» контекстного меню, вызываемого правым щелчком по ячейке со звуковыми единицами таблицы элементов транскрипции (рис. 4.5.5.1), двойным щелчком по этой ячейке, либо щелчком по ссылке «след >» справа от ячейки. При этом текущий выбранный для озвучивания данной звуковой единицы кандидат запрещается, подбор звуковых единиц производится заново, а запрещенный кандидат помечается как запрещенный в таблице кандидатов: в столбце «порядок» ему присваивается очередной номер.

Мон...	Основной тон	Д...	Энергия	<<	Зву...	>>	Fix	<<	ID п...	>>	Fix	Порядок	Кан...
d'	186 177	71	2033 1577 3113		195366	след >	<input type="checkbox"/>		242443	след >	<input type="checkbox"/>		239580
e0	169 161 148 138	31	7204 6807 5165		242444		<input type="checkbox"/>		242445	след >	<input type="checkbox"/>		242445
n'	137 128	126	3112 1318 216		242446		<input type="checkbox"/>		2		Запретить этот юнит		

Мон...	Основной тон	Д...	Энергия	<<	Зву...	>>	Fix	<<	ID п...	>>	Fix	Порядок	Кан...
d'	186 177	71	2033 1577 3113		195366	след >	<input type="checkbox"/>		242443	след >	<input type="checkbox"/>		239580
e0	169 161 148 138	31	7204 6807 5165		242444		<input type="checkbox"/>		242445	след >	<input type="checkbox"/>		242445
n'	137 128	126	3112 1318 216		242446		<input type="checkbox"/>		256306	след >	<input type="checkbox"/>		243639

Мон...	Основной тон	Д...	Энергия	<<	Зву...	>>	Fix	<<	ID п...	>>	Fix	Порядок	Кан...
d'	186 177	71	2033 1577 3113		195366	след >	<input type="checkbox"/>		142212	след >	<input type="checkbox"/>		419747
e0	169 161 148 138	31	7204 6807 5165		142213		<input type="checkbox"/>	< пред	92371	след >	<input type="checkbox"/>		423727
n'	137 128	126	3112 1318 216		92372		<input type="checkbox"/>		92373	след >	<input type="checkbox"/>	0	242445

Рис. 4.5.5.1. Запрет кандидата.

Отмена запрета выбора последнего запрещенного кандидата производится путем щелчка по ссылке «< пред» слева от ячейки со звуковой единицей (рис. 4.5.5.2).

Мон...	Основной тон	Д...	Энергия	<<	Зву...	>>	Fix	<<	ID п...	>>	Fix	Порядок	Кан...
d'	186 177	71	2033 1577 3113		195366	след >	<input type="checkbox"/>		142212	след >	<input type="checkbox"/>		90161
e0	169 161 148 138	31	7204 6807 5165		142213		<input type="checkbox"/>	< пред	92371	след >	<input type="checkbox"/>		92371
n'	137 128	126	3112 1318 216		92372		<input type="checkbox"/>		92373	след >	<input type="checkbox"/>		92554

Рис. 4.5.5.2. Отмена запрета кандидата.

Отмена запрета выбора всех ранее запрещенных кандидатов производится путем выбора пункта «Отменить все запреты» контекстного меню, вызываемого правым щелчком по ячейке со звуковыми единицами (рис. 4.5.5.3).

Мон...	Основной тон	Д...	Энергия	<<	Зву...	>>	Fix	<<	ID п...	>>	Fix	Порядок	Кан...
d'	186 177	71	2033 1577 3113		195366	след >	<input type="checkbox"/>		142212	след >	<input type="checkbox"/>		90161
e0	169 161 148 138	31	7204 6807 5165		142213		<input type="checkbox"/>	< пред	92371	след >	<input type="checkbox"/>		92371
n'	137 128	126	3112 1318 216		92372		<input type="checkbox"/>				<input type="checkbox"/>		

Запретить этот юнит
 Отменить все запреты

Рис. 4.5.5.3. Отмена запрета всех кандидатов.

Имеется возможность зафиксировать выбор кандидата для озвучивания монофона путем включения соответствующей галочки в столбце «Fix» таблицы элементов транскрипции (рис. 4.5.5.4). При этом для озвучивания соответствующего монофона будет использоваться выбранный кандидат. При запрете зафиксированного кандидата фиксация его выбора снимается.

Мон...	Основной тон	Д...	Энергия	<<	Зву...	>>	Fix	<<	ID п...	>>	Fix	▲	Порядок	Кан...
d'	186 177	71	2033 1577 3113		195366	след >	<input type="checkbox"/>		142212	след >	<input type="checkbox"/>			90161
e0	169 161 148 138	31	7204 6807 5165		142213		<input type="checkbox"/>	< пред	92371	след >	<input checked="" type="checkbox"/>			92371
n'	137 128	126	3112 1318 216		92372		<input type="checkbox"/>		92373	след >	<input type="checkbox"/>			92554

Рис. 4.5.5.4. Фиксация выбранного кандидата.

Двойным щелчком по строке таблицы кандидатов можно выбрать конкретный кандидат для озвучивания монофона, отображаемого в текущей строке таблицы элементов транскрипции (рис. 4.5.5.5). После этого данный кандидат будет зафиксирован для озвучивания соответствующего монофона.

Мон...	Основной тон	Д...	Энергия	<<	Зву...	>>	Fix	<<	ID п...	>>	Fix	▲	Порядок	Кан...
d'	186 177	71	2033 1577 3113		374999	след >	<input type="checkbox"/>		123117	след >	<input type="checkbox"/>			90161
e0	169 161 148 138	31	7204 6807 5165		123118		<input type="checkbox"/>	< пред	90161	след >	<input checked="" type="checkbox"/>			92371
n'	137 128	126	3112 1318 216		90162		<input type="checkbox"/>		201814	след >	<input type="checkbox"/>			92554

Рис. 4.5.5.5. Выбор кандидата из таблицы.

Запрещенные кандидаты выделяются в таблице кандидатов темно-красным цветом, однако их также можно выбрать и зафиксировать для озвучивания монофона. Выбор фиксированного кандидата имеет больший приоритет, чем его запрет: если зафиксированный кандидат был ранее запрещен, он тем не менее будет выбран.

Поскольку таблица кандидатов позволяет сортировать кандидаты по значениям признаков (основной тон, длительность, энергия), то таким образом можно выбирать наиболее подходящий кандидат по одному из этих признаков.

Сортировка кандидатов по любому из признаков производится путем щелчка по заголовку соответствующего столбца таблицы кандидатов (рис. 4.5.5.6). Повторный щелчок по заголовку меняет порядок сортировки (по возрастанию / убыванию).

Кандидаты (289 шт.)				
Порядок	Кандидаты (ID)	Основной тон	Длительность	Энергия
	85220	165	72	8146
	86306	273	97	8020
	86486	270	69	8062
	89837	182	60	7831
	90161	294	57	8140
	92371	134	127	7889
	92554	140	86	7932
	93635	277	64	7917
	96519	156	64	8180

Рис. 4.5.5.6. Сортировка кандидатов.

Если в использовании признаков кандидатов (основной тон, длительность, энергия) нет необходимости, то их можно отключить путем включения галочки «Расширенная информация по кандидатам» (DB) в окне конфигурации модуля (см. разд. 4.4.3.2. Диалог конфигурации модуля). При этом соответствующие столбцы с признаками в таблице кандидатов будут скрыты (рис. 4.5.5.7).

Кандидаты (289 шт.)	
Порядок	Кандидаты (ID)
	85220
	86306
	86486
	89837
	90161
	92371
	92554
	93635
	96519

Рис. 4.5.5.7. Отключение отображения признаков кандидатов.

После любой из операций, влияющих на подбор звуковых единиц, производится либо только подбор звуковых единиц для озвучивания фразы, либо подбор звуковых единиц вместе с произнесением фразы – в зависимости от состояния галочки «Произносить фразу при каждом изменении» (см. разд. 4.4.3.2. Диалог конфигурации модуля).

4.5.6. История изменений (Undo)

Любые изменения, влияющие на произношение фразы, отражаются в истории изменений этой фразы. Наличие истории изменений позволяет производить откат любого изменения, примененного к выбранной фразе, кроме удаления фразы (об удалении фраз см. разд. 4.5.2. Создание и редактирование фразы). История изменений хранится отдельно для каждой фразы.

Откат назад по истории изменений фразы производится путем нажатия кнопки «Отменить сделанную операцию» (Undo) (рис. 4.5.6.1). Происходит возврат к предыдущему произношению выбранной фразы.

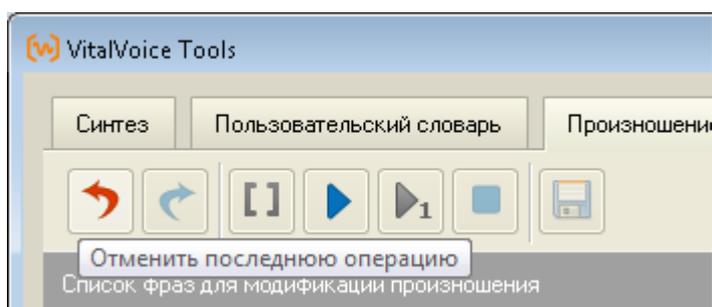


Рис. 4.5.6.1. Отмена операции (Undo).

Продвижение вперед по истории изменений фразы производится путем нажатия кнопки «Вернуть сделанную операцию» (Redo) (рис. 4.5.6.2). Происходит возврат к последнему отмененному изменению произношения выбранной фразы.

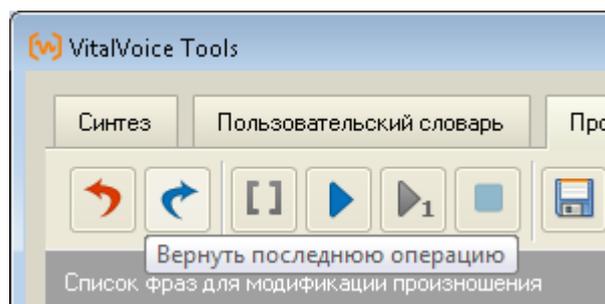


Рис. 4.5.6.2. Возврат отмены операции (Redo).

При перезапуске приложения производится очистка истории изменений всех фраз, в БД ПП история изменений не сохраняется.

4.5.7. Произнесение фраз (Play)

Произнесение фразы с использованием настроенного пользователем произношения производится путем нажатия кнопки «Произнести фразу» (Play) панели инструментов либо путем выбора пункта «Произнести фразу»

контекстного меню, вызываемого правым щелчком по тексту фразы в ячейке таблицы фраз в столбце «Фраза» (рис. 4.5.7.1).

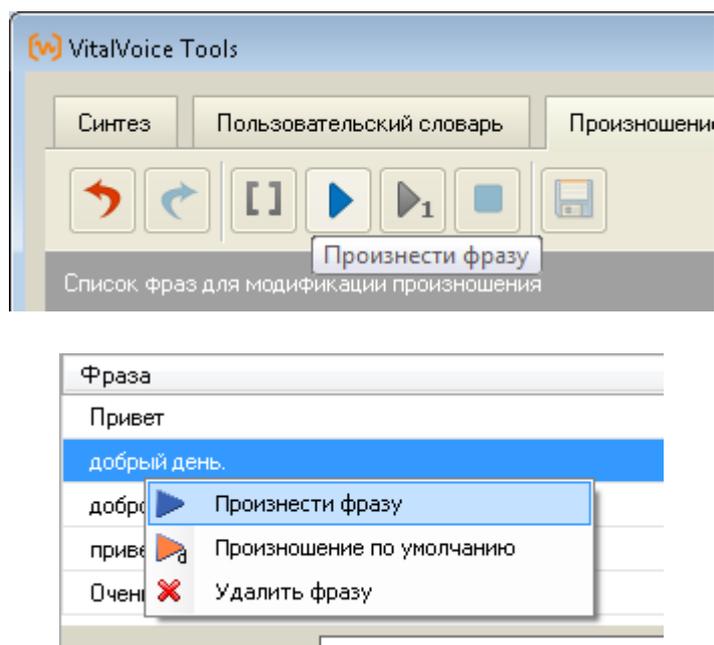


Рис. 4.5.7.1. Произнести фразу (варианты).

Произнесение фразы с использованием произношения по умолчанию (без использования настроенного пользователем произношения) производится путем нажатия кнопки «Произношение по умолчанию» (Play1) панели инструментов, либо путем выбора пункта «Произношение по умолчанию» контекстного меню, вызываемого правым щелчком по тексту фразы в ячейке таблицы фраз в столбце «Фраза» (рис. 4.5.7.2).

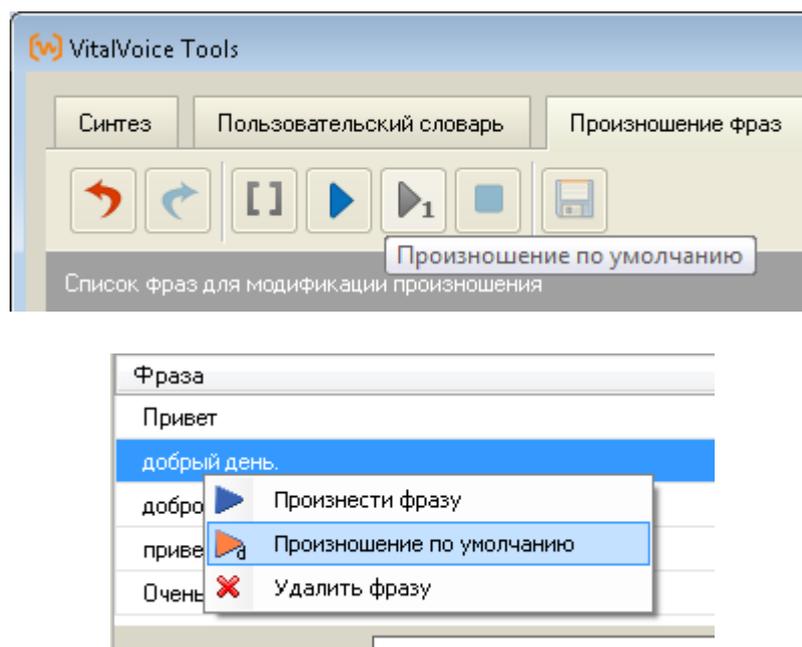


Рис. 4.5.7.2. Произношение по умолчанию (варианты).

Произношение по умолчанию может быть использовано для сравнения с настроенным произношением.

4.5.8. Прочие операции

Для детального анализа звука – например, для обнаружения призвуков, возникающих на стыках звуковых единиц, или для поиска неправильно подобранных звуковых единиц – модуль ПФ сохраняет в .wav файл последнюю произнесенную фразу. Этот файл называется output.wav и находится в папке `ApplicationData\Speech Technology Center\VVTools`

где

`ApplicationData = C:\Documents and Settings\<user_name>\Application Data`

(указано обычное расположение этой папки для Windows XP).

4.5.9. Пример работы с фразами

В этом разделе приводятся практические советы по нескольким возможным сценариям настройки произношения фраз. Подробное описание возможных операций см. в соответствующих разделах. **Внимание!** Для фиксации произношения фразы «навсегда» (т. е. при изменении версий движка и голоса) для данного голоса нужно **зафиксировать все аллофоны-кандидаты нажатием на заголовок столбца Fix**. Желательно также прослушивать получаемое звучание фразы после каждого изменения (по умолчанию опция автоматического проигрывания включена).

4.5.9.1. Добавление и удаление паузы

При желании после любого слова в предложении может быть добавлена или удалена пауза. Это может потребоваться при неудачном стыке звуков на границе слов (например, «в конце_августа»: в этом случае нужна минимальная пауза – «spelling»), при ошибочной работе модуля расстановки пауз или же просто при желании разбить предложение на несколько интонационных фраз. Пауза обозначается символом «#» в таблице элементов транскрипции в левой нижней части окна.

Для добавления паузы следует щёлкнуть правой кнопкой мыши по последней букве слова, после которого нужно добавить паузу, и выбрать её длительность от минимальной «spelling» до максимальной «long»:

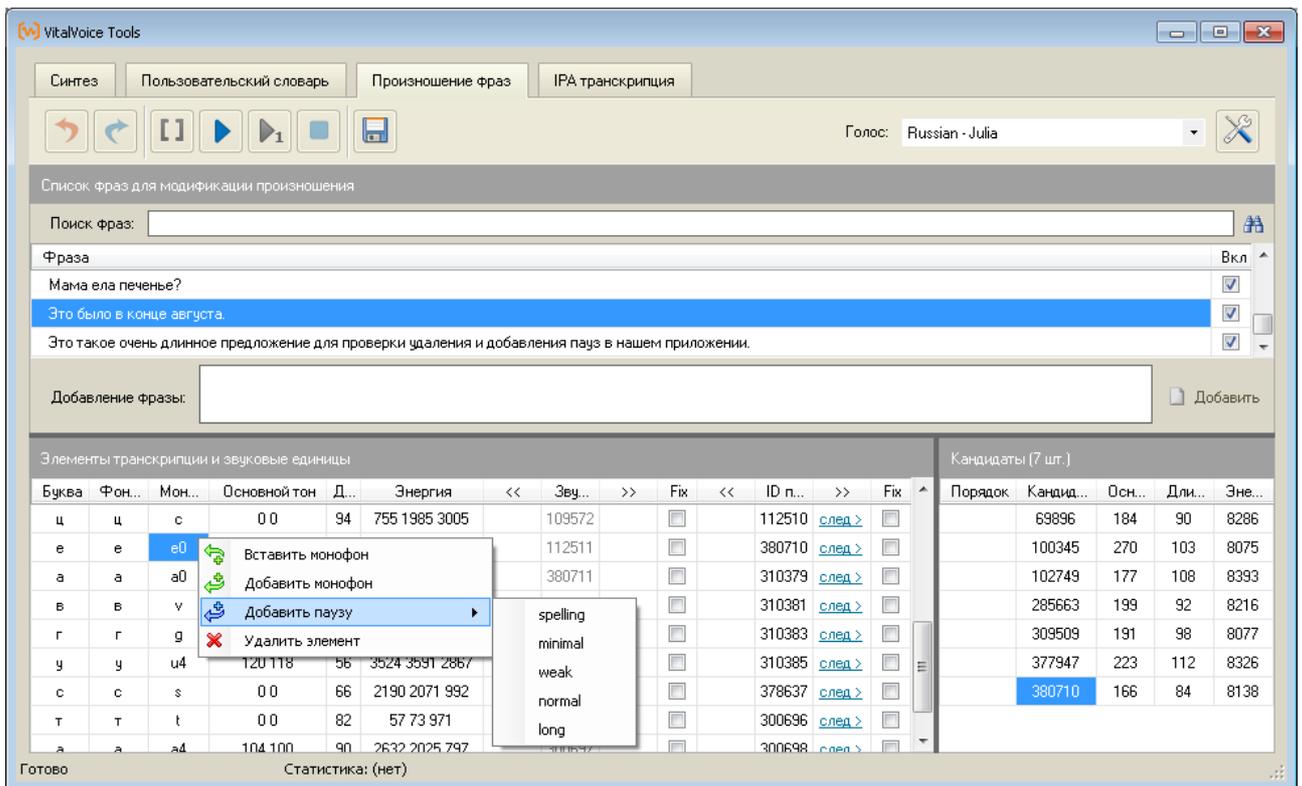


Рис. 4.5.9.1.1. Добавление паузы

Для удаления паузы следует щёлкнуть правой кнопкой по значку «#» паузы, которую вы хотите удалить.

4.5.9.2. Выделение отдельного слова

Для выделения важного слова или просто более чёткого и медленного его произнесения можно добавить небольшие паузы перед ним и/или после него, а также (после добавления пауз, если планируется его делать) подобрать более длинные, громкие или выделенные по высоте единицы.

Для этого нужно выбрать в таблице элементов транскрипции строку, соответствующую нужному звуку (стоит начать с ударного гласного слова). При этом в таблице кандидатов справа появятся возможные аллофоны-кандидаты с указанием их основного тона (высоты звука, для глухих = 0), длительности и энергии (громкости). Кандидаты можно отсортировать по возрастанию или убыванию значений в любом из этих столбцов. Имеет смысл начинать настройку с выбора кандидатов большей длительности. При выборе конкретных кандидатов из таблицы справа изменённые аллофоны фиксируются автоматически (для правой половины полуфонов – второй столбец Fix).

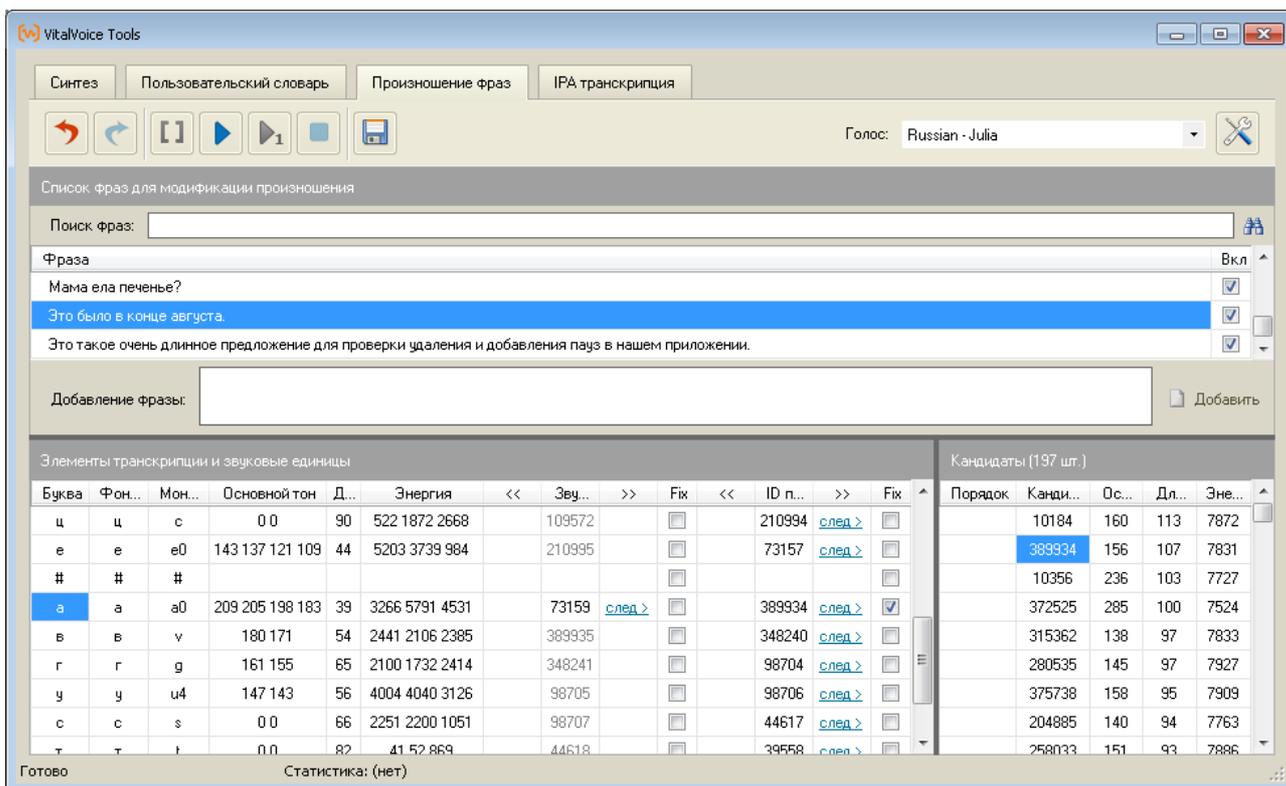


Рис. 4.5.9.2.1. Выбор аллофона для выделения слова «августа» во фразе «Это было в конце августа»

4.5.9.3. Изменение интонации

Иногда из-за особенностей работы алгоритма подбора элементов для синтеза предложение с вопросом может быть прочитано как повествовательное и т. п. В этом случае можно вручную настроить произнесение, выбрав кандидаты с другой частотой основного тона (высотой). Для усиления вопроса следует выбрать для ударного гласного (а также, возможно, и для других аллофонов в ударном слоге) слова, на котором должен звучать вопрос, кандидат с большим значением основного тона. Для завершающей интонации следует наоборот выбрать кандидат с более низким значением. При этом аллофоны должны оставаться достаточно длительными.

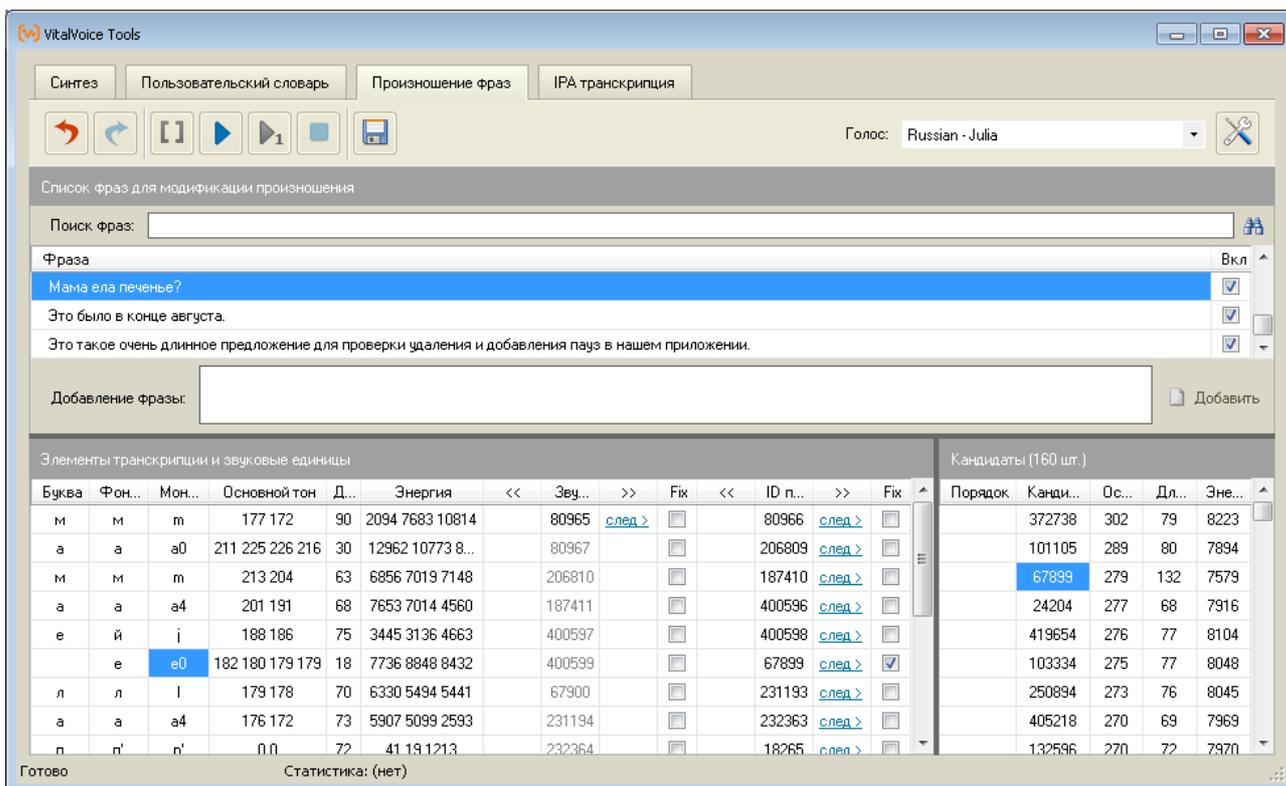


Рис. 4.5.9.3.1. Перенос вопроса на слово «ела» в предложении «Мама ела печенье?». При этом аллофон ударного гласного в слове «печенье» заменяется кандидатом с меньшим значением осинового тона

4.5.9.4. Запрет неудачного аллофона

В случае неудачного прочтения какого-либо звука или нескольких звуков в предложении («кваканье», «бульканье», щелчки и т. п.) можно запретить неудачный звук, а также подобрать наиболее удачный по звучанию вариант. Для этого в строке, соответствующей нужному звуку, нужно нажать ссылку «след >» в столбце «<>>» после столбца «ID правого полуфона» (для аллофонов после паузы также можно выбирать варианты левой половины в первом столбце «<>>»). При этом будет выбран следующий аллофон, наиболее подходящий с точки зрения алгоритма подбора элементов, а предыдущий будет автоматически запрещён. При нажатии на ссылку «< пред» будет возвращён предыдущий вариант, и запрет отменится. При последовательном нажатии на «след >» будет осуществляться перебор всех возможных кандидатов из таблицы справа. При необходимости можно сразу же в правой таблице выбрать кандидат с нужными характеристиками основного тона, длительности и энергии.

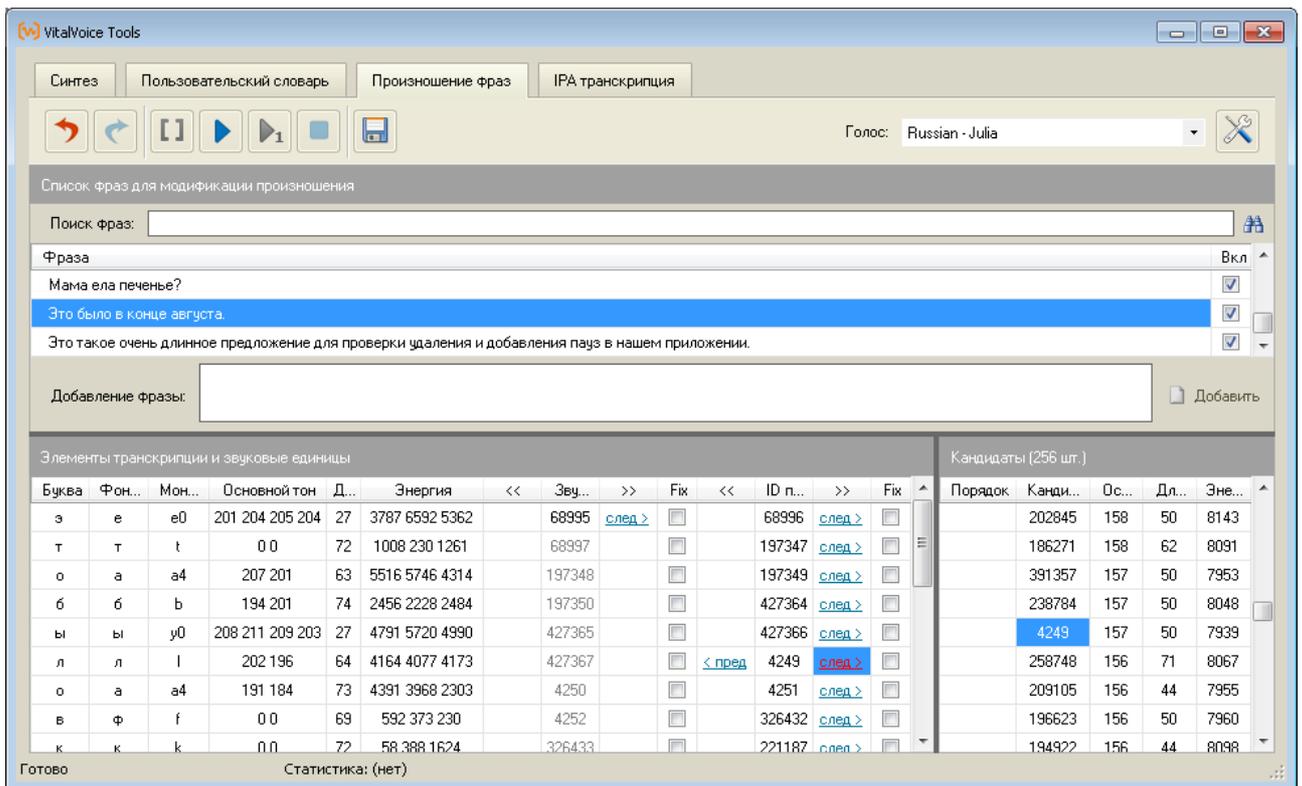


Рис. 4.5.9.4.1. Замена аллофона «л» в слове «было» другим кандидатом

4.6. Дополнительная информация

4.6.1. Алгоритм Unit Selection

Подбор звуковых элементов осуществляется методом Unit Selection, при котором из голосовой базы, содержащей около 10 часов речи диктора, выбираются фрагменты (отдельные дифоны, полуфоны или более крупные фрагменты) которые затем соединяются вместе, образуя требуемую фразу.

4.6.2. Особенности подбора звуковых элементов

Метод Unit Selection позволяет синтезировать речь очень натурального качества, близкую к речи реального диктора. Однако, одна из особенностей алгоритма Unit Selection подбора звуковых элементов состоит в том, что звучание некоторых фраз (каких именно – заранее предсказать невозможно) может показаться не таким, как ожидает слушатель. Это обусловлено тем, что, во-первых, формальные признаки, по которым осуществляется подбор элементов, не всегда полно описывают звучание с точки зрения слухового восприятия, и, во-вторых, в голосовой базе может не найтись элемента с требуемыми параметрами. В этом случае вместо него алгоритм может подобрать элемент, по формальным признакам близкий к требуемому, который, появляясь в составе синтезируемой фразы, может «резать слух».

Для тех фраз, сгенерированных синтезатором, произношение которых отличается от ожидаемого, оно может быть модифицировано вручную. Для этого модуль ПФ позволяет запретить выбор некоторых звуковых единиц, чтобы исключить нежелательный звук, тогда вместо запрещенной звуковой единицы (кандидата) будет выбран другой. Возможно, при этом будут также выбраны другие кандидаты вместо соседних или близко расположенных элементов во фразе, либо могут быть оставлены те же, по желанию пользователя. Настройку произношения фразы можно выполнять до тех пор, пока оно не будет полностью устраивать пользователя. Ограничением может служить только исчерпание всех возможных кандидатов для выбранного звука.

4.6.2.1. Особенности дифонного синтеза

Vital Voice TTS в качестве звуковой единицы использует дифон. В отличие от аллофона, границы которых совпадают с границами гласных или согласных звуков, дифон представляет собой такой звуковой элемент, начало которого находится примерно посередине одного гласного или согласного звука, а конец – примерно посередине следующего. На *рис. 4.6.2.1.1* в верхнем ряду показано деление звука на аллофоны, в нижнем ряду – на дифоны.

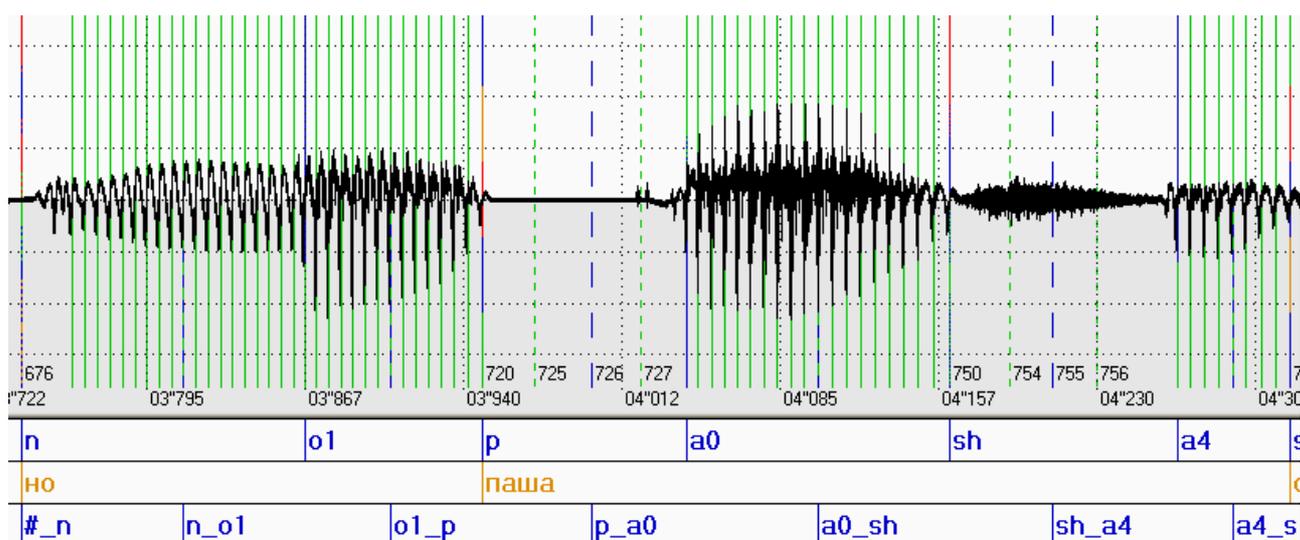


Рис. 4.6.2.1.1. Аллофоны и дифоны.

Таким образом, каждый дифон состоит из двух частей. Так, в примере, приведенном на *рис. 4.6.2.1.1*, гласному звуку «a0» в слове «Паша» соответствует правая половинка дифона «p_a0» и левая половинка дифона «a0_sh». Такие половинки, у которых одна из границ совпадает с границей аллофона, а другая – с границей дифона, называются полуфонами.

В отличие от аллофонов, стыковка дифонов происходит на середине гласных или согласных звуков. В тех случаях, когда предыдущий или следующий гласный или согласный звук отсутствует, – например, для первого элемента

после паузы или последнего перед паузой, – вместо дифона берется полуфон. Пример полуфона «#_n» показан на *рис. 4.6.2.1.1*.

Также отдельные полуфоны могут быть задействованы при сборке звука в тех случаях, когда целый дифон с требуемыми характеристиками не может быть найден в звуковой базе, либо если кандидатов–дифонов в базе недостаточно. В таких случаях недостающие дифоны собираются из двух половинок, взятых из разных участков звуковой базы. Подобное случается редко.

Для возможности сборки дифонов из полуфонов каждый полуфон звуковой базы имеет свой идентификатор, по которому этот полуфон может быть найден в базе. Дифон как целый звуковой элемент, состоящий из двух полуфонов, обозначается идентификатором (ID) своего первого полуфона.

В большинстве случаев дифон представляет собой два полуфона, идущих подряд, при этом ID составляющих полуфонов также идут подряд и отличаются на единицу. В тех редких случаях, когда дифон собирается из двух половинок, взятых из разных участков звуковой базы, ID составляющих его полуфонов идут не подряд.

На *рис. 4.6.2.1.2* показана транскрипция и разложение на звуковые единицы фразы «Но Вася занят».

Элементы транскрипции и звуковые единицы													Кандидаты (89 шт.)					
Буква	Фон...	Мон...	Основной тон	Д...	Энергия	<<	Зву...	>>	Фик	<<	ID п...	>>	Фик	Порядок	Канди...	Ос...	Дл...	Эне...
н	н	n	187 179	102	1433 4709 6294		332096	след >	<input type="checkbox"/>		332097	след >	<input type="checkbox"/>		407430	192	88	7853
о	о	o1	187 200	95	7853 7653 5846		332098		<input type="checkbox"/>		332099	след >	<input type="checkbox"/>		408384	137	110	7405
в	в	v	205 208	61	3345 3340 5364		332100		<input type="checkbox"/>		428304	след >	<input type="checkbox"/>		416294	244	81	7514
а	а	a0	208 204 202 198	24	6140 5573 4573		428305		<input type="checkbox"/>		428306	след >	<input type="checkbox"/>		425017	152	72	7913
с	с'	s'	0 0	91	2078 1830 1845		428307		<input type="checkbox"/>		317195	след >	<input type="checkbox"/>		425149	200	76	7901
я	а	a4	191 191	68	4943 4938 4071		317196		<input type="checkbox"/>		317197	след >	<input type="checkbox"/>		428306	175	112	7807
э	э	z	189 183	71	2795 2191 3210		317198		<input type="checkbox"/>		65783	след >	<input type="checkbox"/>		432797	146	103	7937

Рис. 4.6.2.1.2. Звуковые единицы при настройке произношения.

Элементы 332097, 332099, 428304, 428306, 317195, 317197, 65783 и другие представляют собой дифоны. О том, что это именно дифоны, а не отдельные полуфоны, свидетельствует тот факт, что ID составляющих их полуфонов идут подряд. Так, дифон 428304 «v_a0», состоит из половинок 428304+428305, дифон 428306 «a0_s'» – из половинок 428306+428307. Можно видеть, что вторые полуфоны этих дифонов (для гласных и согласных звуков они являются левыми половинками) подсвечены серым. Они недоступны для редактирования пользователем: для них невозможно запретить или выбрать кандидаты. Причина этого состоит в том, что дифоны рассматриваются и участвуют в синтезе как неразрывные элементы, а не как отдельные полуфоны.

Отдельно друг от друга редактируются только те полуфоны, которые не составляют диффон. Таким полуфоном является, например, элемент 332096 – он представляет собой полуфон «#_n», следующий после паузы.

4.6.3. Изменения алгоритма подбора элементов

Алгоритм подбора звуковых элементов может быть изменен при переходе к очередной версии движка синтеза, поэтому произношение каждой отдельной фразы при установке новой версии движка может измениться. Обычно изменения в алгоритме подбора звуковых элементов приводят к улучшению произношения, однако для отдельных фраз произношение может стать хуже, хотя бы даже по субъективному мнению пользователя – например, фраза может звучать не так, как звучала раньше, в «старом добром» привычном варианте.

Тем более следует учитывать возможность изменения произношения фразы, специально настроенной пользователем с применением модуля ПФ, когда пользователь вручную определяет оптимальное с его точки зрения произношение.

Важно иметь в виду, что оптимальность звучания фразы, обеспеченная модулем ПФ для текущей версии движка синтеза за счет запрета выбора отдельных кандидатов, не гарантирует сохранения такого же звучания в другой версии движка.

Предположим, движок версии 0.3.20.5 для фразы «икс» по умолчанию выбирал звуковые единицы «N1_1, N2_1, N3_1, N4_1». Пусть пользователь запретил выбор кандидата N1_1, затем запретил N1_2, затем N1_3, и добился оптимального с его точки зрения произношения «N1_4, N2_4, N3_4, N4_4».

Движок версии 0.3.22.1 для той же фразы «икс» может по умолчанию выбирать звуковые единицы «N1_5, N2_5, N3_5, N4_5», и произношение фразы может стать иным, чем было после настройки для движка 0.3.20.5, хотя запрет кандидатов N1_1, N1_2, N1_3 для первого элемента останется в силе.

Для обеспечения точного сохранения произношения фразы при переходе к новой версии TTS модуль ПФ позволяет зафиксировать выбранные кандидаты для всех звуковых единиц фразы, нажав кнопку «Зафиксировать произношение фразы» (галочка) (см. разд. 4.4.3.4. *Таблица элементов транскрипции*). В таком случае движок будет подбирать для озвучивания каждого элемента данной фразы соответствующий фиксированный кандидат, и произношение фразы с большой вероятностью (за исключением случаев изменения формата голосовой базы) будет точно совпадать с однажды настроенным.

5. IPA-ТРАНСКРИПЦИЯ

5.1. Назначение

Модуль редактирования словаря IPA-транскрипций предназначен для создания и редактирования транскрипций для отдельных слов. Этот словарь используется движком синтеза для модификации произношения отдельных слов. Модуль позволяет для любого слова задать его транскрипцию, записав ее в виде последовательности символов международного фонетического алфавита IPA.

5.2. Хранение данных

IPA-транскрипции слов хранятся в пользовательском словаре IPA-транскрипций (далее – IPA-словарь). Транскрипция, созданная для конкретного слова, не зависит от голоса диктора (она не содержит информации об элементах звуковой базы), однако при синтезе она может использоваться как для всех существующих голосов, так и для одного конкретного голоса, согласно тому, как определит пользователь для данного слова.

IPA-словарь, создаваемый пользователем при помощи модуля редактирования словаря IPA-транскрипций VVTools, хранится в файле `Transform.Lexicon.txt`, который находится в папке с пользовательскими данными, путь к которой определяется ключом реестра `HKLM\SOFTWARE\Speech Technology Center\TTS\Languages\<...>\UserDataPath` где `<...>` = Russian для русского языка. При отсутствии данного ключа реестра IPA-словарь будет сохраняться в текущей директории.

5.3. Использование совместно с движком синтеза (TTS Engine)

Подключение произвольного IPA-словаря производится путем подачи `ssml`-тега `lexicon` на вход движка синтеза. При этом в атрибуте `uri` этого тега указывается путь к файлу IPA-словаря: `<lexicon uri="c:\path\lexicon.txt"/>`.

IPA-словарь `Transform.Lexicon.txt`, находящийся в папке с пользовательскими данными (см. разд. 5.2. Хранение данных) может быть подключен в качестве IPA-словаря, используемого по умолчанию, без использования `ssml`-тега `lexicon`. Для этого в движке синтеза требуется задать ключ конфигурации `("TTS/phonetic/transform/lexicon_default", "on")`

После подключения IPA-словаря для каждого слова, подаваемого на вход движка, вместо произношения, создаваемого движком по стандартным

правилам транскрипции русского языка, будет подставлено произношение из IPA-словаря, соответствующее этому слову, при наличии такового в словаре.

5.4. Интерфейс и использование приложения

5.4.1. Необходимые условия для работы модуля

Для работы модуля наличие установленного движка синтеза (TTS Engine) не требуется. Редактирование IPA-словаря возможно без установленного движка. Однако, произнесение слов и транскрипций без установки движка недоступно.

5.4.2. Запуск модуля

Модуль запускается при первом переходе на соответствующую вкладку в главном окне приложения. При запуске модуля происходит загрузка IPA-словаря из папки с пользовательскими данными (см. разд. 5.2. *Хранение данных*).

5.4.3. Интерфейс модуля

Внешний вид интерфейса модуля изображен на *рис. 5.4.3*.

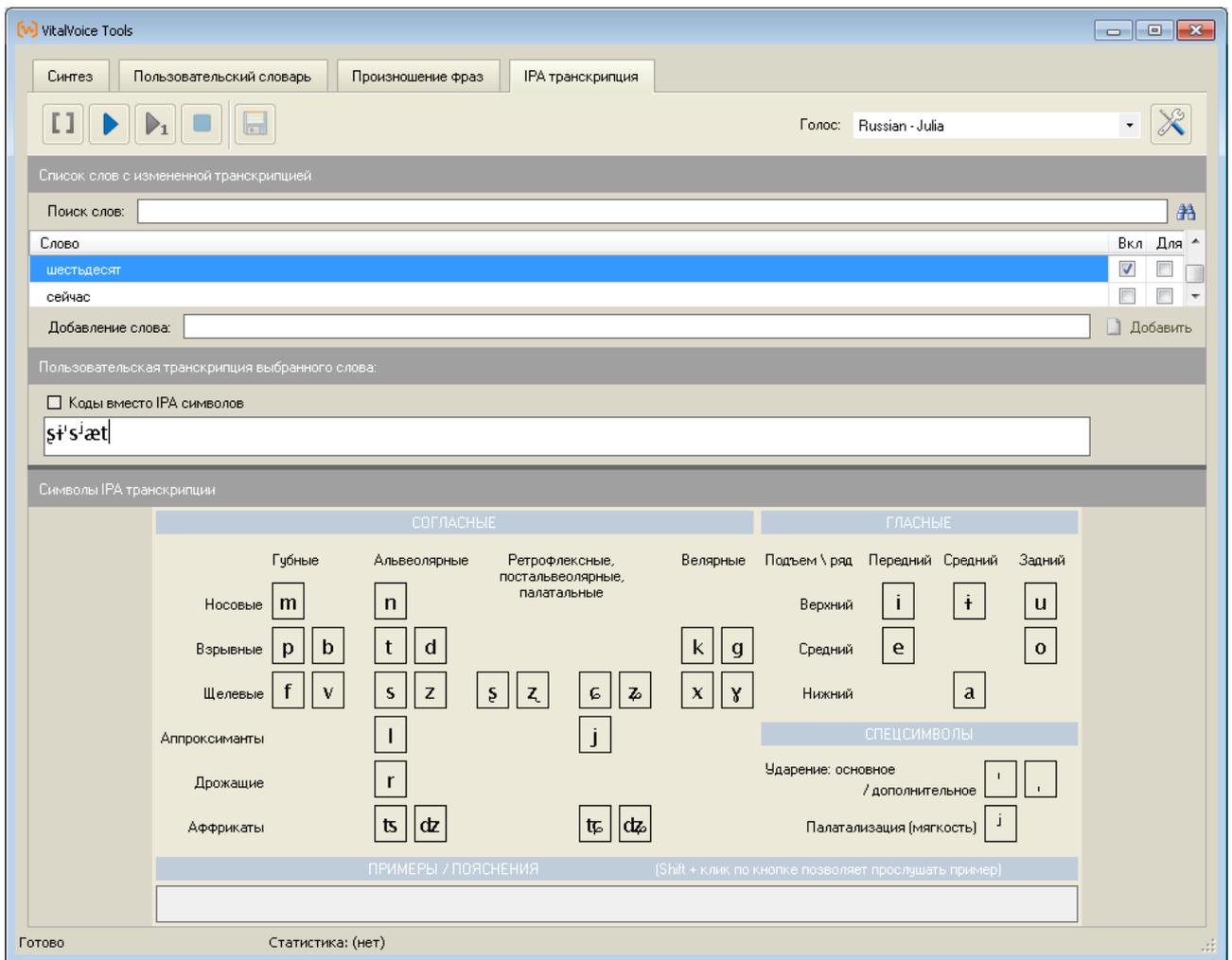


Рис. 5.4.3. Внешний вид интерфейса модуля «Транскриптор».

5.4.3.1. Панель инструментов

Панель инструментов расположена в верхней части окна приложения, снизу от элемента управления для выбора модулей. Кнопки на панели инструментов позволяют производить следующие операции:

Генерация транскрипции ([]). Создает для слова транскрипцию по умолчанию при помощи движка синтеза. *Подробнее о редактировании транскрипции см. разд. 5.5.3. Модификация транскрипции.*

Произнести слово (Play). Произносит слово с применением созданной транскрипции. *Подробнее о редактировании транскрипции см. разд. 5.5.3. Модификация транскрипции.*

Произношение по умолчанию (Play1). Произносит слово с применением произношения по умолчанию (транскрипция по стандартным правилам русского языка).

Стоп (Stop). Прервать произнесение слова.

Сохранить (Save). Сохраняет файл IPA-словаря. *Имя и расположение файла описаны в разд. 5.2. Хранение данных.*

Голос. Выпадающий список выбора голоса позволяет выбрать голос для озвучивания слов и транскрипций. Выбранный пункт определяет голос, которым будет озвучена транскрипция в интерфейсе модуля, но как список слов, так и транскрипции, созданные для каждого слова, не зависят от выбранного голоса.

Настройки (Tools). Открывает диалоговое окно конфигурации модуля (рис. 5.4.3.1)

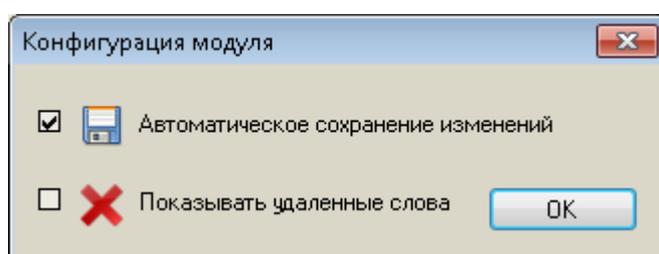


Рис. 5.4.3.1. Диалог конфигурации модуля.

5.4.3.2. Диалог конфигурации модуля

Диалоговое окно конфигурации предназначено для изменения параметров конфигурации модуля. Возможно изменение следующих параметров.

Автоматическое сохранение изменений: включает режим автоматического сохранения в IPA-словарь всех вносимых изменений.

Показывать удаленные слова: включает отображение слов, удаленных из IPA-словаря (при сохранении в файле такие слова закоментированы специальным образом).

5.4.3.3. Список слов

Табличное представление «Список слов с измененной транскрипцией» в верхней части окна под панелью инструментов содержит список слов, содержащихся в IPA-словаре. Добавление слова производится в отдельной строке, см. разд. 5.4.3.4. *Строка добавления слова.*

Ячейки столбца «Слово» содержат текстовое представление слова. Редактирование слова производится прямо в ячейке.

Ячейки столбца «Вкл» позволяют включать / отключать использование слова в движке синтеза.

Ячейки столбца «Для» позволяют включать / отключать привязку слова к конкретному голосу в движке синтеза. Слово привязывается к тому голосу, который выбран в панели инструментов. Слова, не привязанные к голосу, используются для всех голосов. Слова, привязанные к другому голосу (не к тому, который выбран в панели инструментов), не отображаются в списке.

5.4.3.4. Строка добавления слова

Строка для добавления нового слова, находящаяся под списком слов, позволяет добавлять новое слово в IPA-словарь. Кнопка справа от этой строки непосредственно осуществляет добавление слова в список, при этом для введенного слова создается транскрипция по умолчанию. Для вновь создаваемого слова привязка к голосу отсутствует.

5.4.3.5. Фильтр слов

Строка для ввода текста «Фильтр слов», находящаяся над списком слов, позволяет фильтровать слова по содержимому.

При вводе в строку каких-либо символов в списке слов отображаются только те слова, которые содержат в виде подстроки введенные пользователем символы. Удобно при большом количестве слов в словаре.

5.4.3.6. Пользовательская транскрипция слова

Строка для ввода пользовательской транскрипции слова позволяет задавать транскрипцию для выбранного слова из списка. Транскрипция задается в виде последовательности IPA символов, которые могут быть введены при помощи виртуальной клавиатуры (см. разд. 5.4.3.7. *Виртуальная IPA-клавиатура*).

Кнопка «Сохранить» непосредственно осуществляет обновление транскрипции для выбранного слова. В использовании этой кнопки нет необходимости, поскольку каждое внесенное изменение в строке ввода пользовательской транскрипции сохраняется автоматически.

5.4.3.7. Виртуальная IPA-клавиатура

В нижней части окна приложения находится виртуальная IPA-клавиатура, позволяющая вводить символы IPA-транскрипции в строку для ввода пользовательской транскрипции (см. разд. 5.4.3.6. *Пользовательская*

транскрипция слова). Щелчок по кнопке виртуальной клавиатуры добавляет соответствующий символ в позицию текстового курсора указанной строки.

5.4.3.8. Представления транскрипции

Над строкой для ввода пользовательской транскрипции слова расположена галочка «Коды вместо IPA символов». Включение этой галочки позволяет отображать транскрипцию в виде кодов, так что она может быть вставлена в виде тегов непосредственно в текст. Транскрипция может быть представлена как в чистом виде (как атрибут тега `phoneme`), так и вместе с этим тегом (рис. 5.4.3.8.1).

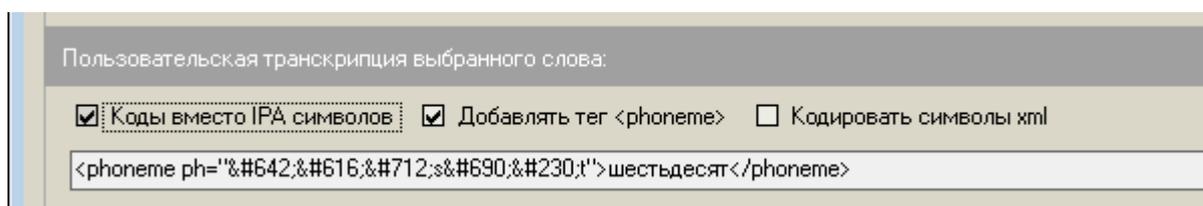


Рис. 5.4.3.8.1. Представление транскрипции в виде кодов.

Имеется также возможность дополнительно кодировать специальные xml символы в виде кодов: `< = <`; `& = &`; и т.п. (рис. 5.4.3.8.2).

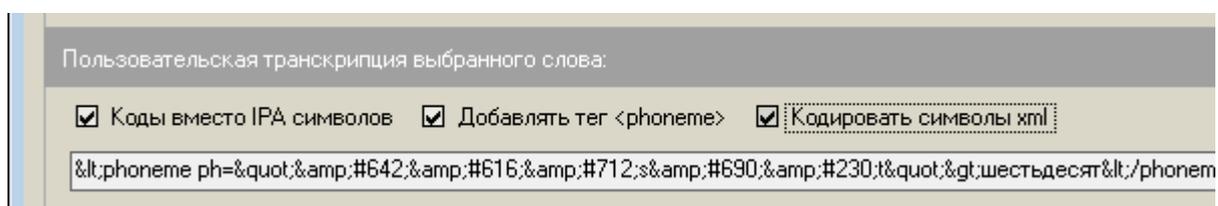


Рис. 5.4.3.8.2. Кодирование специальных символов xml.

5.5. Работа с приложением

5.5.1. Выбор голоса для озвучивания транскрипций

Выбор голоса для озвучивания слов и их транскрипций производится путем выбора требуемого элемента выпадающего списка в выпадающем списке выбора голоса (см. разд. 5.4.3.1. Панель инструментов).

Выбор голоса влияет на озвучивание слов и транскрипций в интерфейсе модуля и на состав элементов, отображаемых в списке. Слова и транскрипции из IPA-словаря, не привязанные к какому-либо голосу, всегда присутствуют в списке; их использование в движке синтеза не зависит от выбранного голоса. Слова, привязанные к конкретному голосу, отображаются в списке только при выборе данного голоса, и в движке используются только совместно с ним.

5.5.2. Создание и редактирование слова

Редактирование слова производится в ячейке таблицы слов (см. разд. 5.4.3.3. *Список слов*) в столбце «Слово» путем нажатия клавиши F2, либо левого щелчка по текущей ячейке, либо двойного щелчка по ячейке. При изменении слова в ячейке изменение его транскрипции не производится.

Добавление нового слова производится путем записи этого слова в строку добавления слова (см. разд. 5.4.3.4. *Строка добавления слова*) и нажатия кнопки «Добавить» справа от этой строки.

Удаление слова производится путем выбора пункта «Удалить слово» контекстного меню, вызываемого правым щелчком по ячейке со словом. В модуле реализовано логическое удаление слова путем пометки его как неиспользуемого, без физического удаления из IPA-словаря. Для восстановления удаленного слова следует включить отображение удаленных слов, поставив соответствующую галочку «Показывать удаленные слова» (Check Delete) в окне настроек (см. разд. 5.4.3.2. *Диалог конфигурации модуля*) и затем заново включить использование удаленного слова. Физическое удаление слова из IPA-словаря возможно путем удаления неиспользуемого слова в режиме отображения удаленных слов. Физическое удаление слова не может быть отменено средствами, предоставляемыми модулем.

Галочка в ячейке столбца «Вкл» позволяет включать / выключать использование соответствующего слова в движке, аналогично логическому удалению слова.

Галочка в ячейке столбца «Для» позволяет включать / отключать привязку соответствующего слова к выбранному голосу.

5.5.3. Модификация транскрипции

Изначально при добавлении нового слова (см. разд. 5.4.3.4. *Строка добавления слова*) для него генерируется транскрипция по умолчанию. Сгенерировать транскрипцию по умолчанию для слова можно и вручную, используя кнопку «Генерация транскрипции» (см. разд. 5.4.3.1. *Панель инструментов*) (рис. 5.5.3.1).

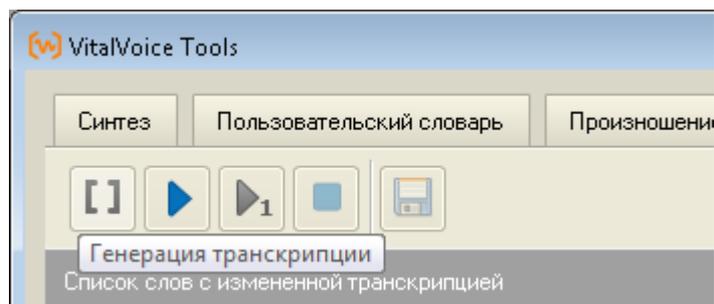


Рис. 5.5.3.1. Генерация транскрипции.

Транскрипцию слова можно модифицировать, редактируя ее в строке для ввода пользовательской транскрипции (см. разд. 5.4.3.6. *Пользовательская транскрипция слова*, 5.4.3.7. *Виртуальная IPA-клавиатура*).

5.5.4. История изменений (Undo)

Пока не реализовано

5.5.5. Произнесение слова (Play)

Произнесение слова с использованием созданной транскрипции производится путем нажатия кнопки «Произнести слово» (Play) панели инструментов (рис. 5.5.5.1).

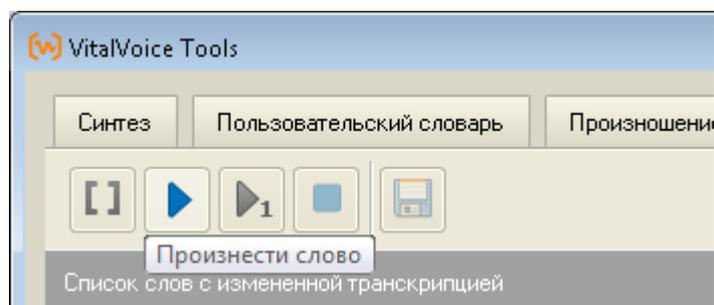


Рис. 5.5.5.1. Произнести слово.

Произнесение слова с использованием транскрипции по умолчанию (по стандартным правилам транскрипции русского языка, без использования созданной пользователем транскрипции) производится путем нажатия кнопки «Произношение по умолчанию» (Play1) панели инструментов (рис. 5.5.5.2).

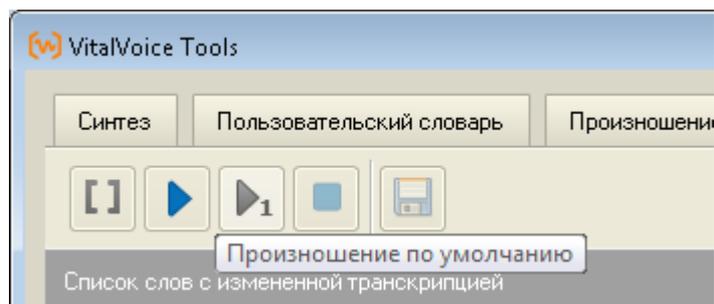


Рис. 5.5.5.2. Произношение по умолчанию.

Произношение по умолчанию может быть использовано для сравнения с произношением созданной пользователем транскрипции.

5.5.6. Сохранение IPA-словаря

Сохранение словаря IPA-транскрипций в файл производится путем нажатия кнопки «Сохранить» панели инструментов (рис. 5.5.6.1).

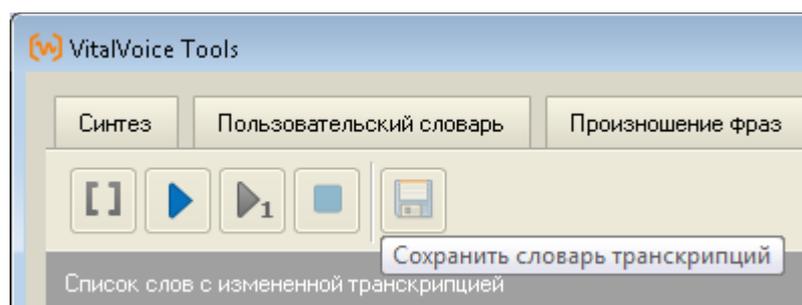


Рис. 5.5.6.1. Сохранение IPA-словаря.

Если в настройках модуля (см. *разд. 5.4.3.2. Диалог конфигурации модуля*) включена галочка «Автоматическое сохранение изменений», то сохранение словаря IPA-транскрипций в файл производится автоматически при любом изменении слова или его транскрипции.

Если пользователь пытается выйти из приложения, а словарь транскрипций не сохранен, модуль выдает соответствующее предупреждение (рис. 5.5.6.2).

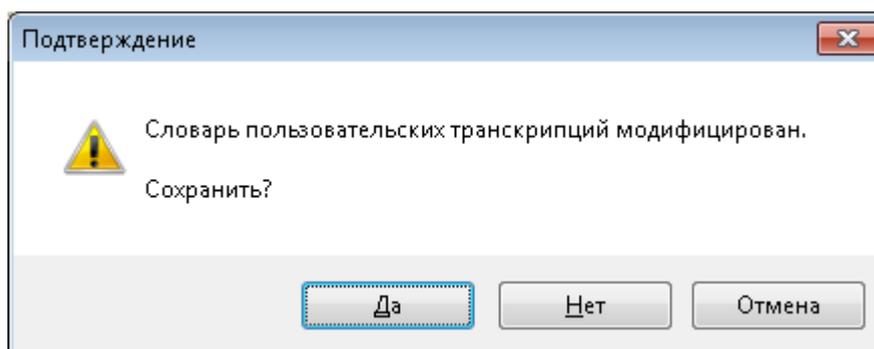


Рис. 5.5.6.2. Предупреждение при выходе.

5.6. Дополнительная информация

5.6.1. IPA транскрипция

Запись транскрипции слов в словаре транскрипций производится с помощью символов алфавита IPA:

http://en.wikipedia.org/wiki/International_Phonetic_Alphabet

5.6.2. Ограничения использования алфавита IPA в движке синтеза

Алфавит IPA предоставляет весьма широкие возможности для записи транскрипций слов с множеством оттенков произношения. С помощью этого алфавита можно записать много различных звуков (все известные фонемы всех языков мира); еще большее разнообразие вариаций произношения может быть записано при помощи диакритик (специальных знаков, надстрочных и подстрочных символов): аспирация, назализация, хриплость, шепот и т.п. – всего более 30 признаков.

При этом в русском языке реализуются только некоторые из всего множества фонем, которые могут быть записаны символами IPA:

http://en.wikipedia.org/wiki/Wikipedia:IPA_for_Russian

Остальные символы IPA не имеют соответствия в фонетике русского языка.

Однако, в целях максимально широкого охвата всех имеющихся символов IPA, чтобы избежать отсутствия звука в том месте, где он, согласно транскрипции, должен быть, в движке синтеза было реализовано преобразование почти всех символов IPA в звуки. Это преобразование основано на экспертном подборе русской фонемы, наиболее близкой к соответствующему IPA-символу по звучанию. Например, фонема [h] используется для озвучивания 11 различных символов IPA. Поэтому, несмотря на то, что каждый из этих 11 символов формально соответствуют разным звукам, все они будут одинаково озвучены движком синтеза в силу банального отсутствия соответствующих звуков в русской фонетике, и, следовательно, в звуковой базе данных, записанной русскоговорящим диктором.

В силу тех же причин движком синтеза не поддерживаются большинство диакритик IPA, означающих различные вариации произношения звуков. Поддерживается только основное ударение и дополнительное ударение для гласных звуков, а также палатализация (мягкость) для согласных звуков.

5.6.3. Особенности обработки IPA транскрипции в движке синтеза

Символы алфавита IPA при обработке их движком синтеза преобразуются в звуки русской фонетики максимально корректно.

Символы, не являющиеся символами IPA, а также некоторые особо экзотические символы IPA игнорируются при обработке движком и не дают звука при произношении.

Для удобства введены некоторые расширения.

Все буквы латинского алфавита, даже те, которые не являются символами IPA, максимально корректно преобразуются в русские фонемы. При этом комбинации символов, вроде sh, zh, ch, sch, и т.п. не поддерживаются, латинские буквы обрабатываются по одной.

Основное ударение может быть обозначено «*» перед слогом или перед гласной, дополнительное ударение – «+» перед слогом или перед гласной, палатализация – «'» после согласной.