

=Speech Auto Tuner= (SAT)

Версия от 2025-04-12

РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ

1. Функциональные характеристики

Программа **Speech Auto Tuner (SAT)** предназначена для автоматической настройки с целью нормализации акустических характеристик сигнала с микрофона в процессе создания речевого аудио контента. Использование средств нормализации позволит стандартизировать уровень речевого сигнала (громкость) и его частотную характеристику (спектр), необходимые для наиболее комфортного и безошибочного слухового восприятия устной речи.

SIA реализован как отдельное **Приложение** для **Windows** и **Linux**.

2. Области применения

Настоящее программное приложение может стать востребованным при подготовке речевого контента во многих Интернет-приложениях, например, в таких как:

- Новостной и развлекательный контент.
- Телеинтервью и беседы.
- Лекции и образовательный контент.
- Вебинары и онлайн-конференции.
- Подкасты, видео-блоги, стриминговые сервисы, социальные сети и многое другое.

Проблемы в восприятии речевого контента, связанные с разбросом акустических параметров микрофона и его расположением, характерны для большинства блогеров, и зачастую приводят к отказу от повторного обращения к ним из-за "бубнящей" (ослабление высоких частот) или "свистящей" (ослабление низких частот) в речи. Программа **Speech Auto Tuner (SAT)** обеспечивает в процессе создания речевого аудио контента автоматическую нормализацию (стандартизацию) акустических характеристик сигнала с микрофона, а именно: уровня громкости и частотного спектра.

3. Начало работы

Стартовое окно SAT, открывающееся после запуска программы, показано на Рис. 1.

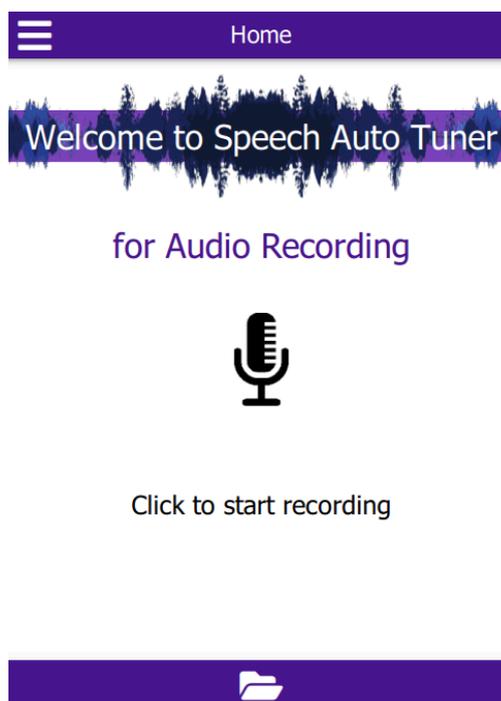


Рис. 1. Стартовое окно SAT

После запуска программы пользователю предоставляются возможность записи нового речевого сигнала с имеющегося у пользователя микрофона, либо вызова из памяти ранее записанного речевого сигнала (PC) от другого источника речи, используя соответствующие значки: , или . В результате записи PC, открывается рабочее окно (рис. 2).

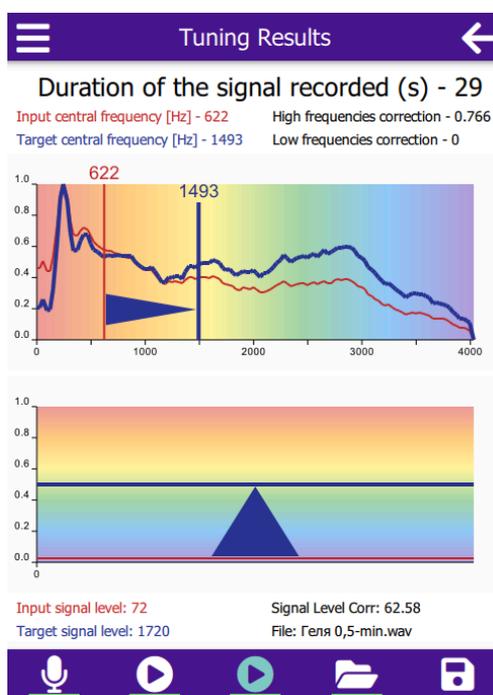


Рис. 2. Рабочее окно SAT

На верхнем графике (рис. 2) отображается АЧС исходного сигнала (красная линия) и нормализованного (синяя линия), а также вертикальные линии (красная и синяя), показывающие положение ЦТ АЧС до и после нормализации спектра. На нижнем графике показаны уровни громкости для исходного РС низкого уровня (красная горизонтальная прямая) и нормализованного РС (синяя линия). Стандартное значение нормализованного уровня РС выбрано равным 0,5 от максимально допустимого уровня РС (см. таблицу 1).

Таблица 1. Значения дБ LUFS для допустимых значений громкости

№	Отн. в дБ LUFS	Отн. в размах	Доли от $2^{16} S(t)$
1	-10	3,16	20739
2	-11	3,55	18461
3	-12	3,98	16466
4	-13	4,47	14661
5	-14	5,01	13081
6	-15	5,62	11661
7	-16	6,31	10386
8	-17	7,08	9256
9	-18	7,94	8254
10	-19	8,91	7355
11	-20	10,0	6554

Duration of the signal recorded (s) - 29

Input central frequency [Hz] - 957

High frequencies correction - 0.508

Target central frequency [Hz] - 1408

Low frequencies correction - 0

В верхней части рабочего окна размещены сведения о длительности РС и используемые данные, необходимые для коррекции его частотного спектра.

Input signal level: 4840

Signal Level Corr: 0.44

Target signal level: 2130

File: Speech Level-Max.wav

В нижней части рабочего окна размещены данные, необходимые для коррекции уровня громкости РС, а также название звукового файла, использовавшегося при тестировании.

После окончания произнесения очередного фрагмента РС через микрофон , осуществляется его запись в папку **data/records** в виде цифрового сигнала с частотой дискретизации 8 кГц с указанием даты и времени начала записи. Первый слева значок  используется для контрольного прослушивания записанного РС, второй – позволяет прослушать нормированный РС. Нормализация АЧС РС и уровня громкости осуществляется автоматически после окончания записи очередного речевого фрагмента. После активации второго значка  осуществляется его запись в папку **data/records** с добавлением имени - **Corrected**.

 14.04.2025.11.38.18 - Corrected

 14.04.2025.11.38.18

В папке **data/tests в качестве тестовых** размещены примеры записей РС с различными уровнем громкости и спектральной характеристики.

4. Вспомогательные значки

В верхней полоске слева рабочего окна находится значок , при нажатии на который открывается дополнительное информационное окно (см. рис. 3),



Рис. 3. Информационное окно

Значок настроек  Settings позволяет изменить значения параметров программы. Внутренняя информация, вызываемая пометкой в правом верхнем квадрате Show advanced, предназначена только для разработчика. Изменения её содержимого пользователем крайне нежелательно и может привести к неправильной работе программы.

При выборе значка:  Results History вызывается содержимое файла “**results**”, в котором сохранены данные РС и его коррекции. Такие как: имя файла, время записи сигнала, его длительность, значения частоты и уровня исходного и скорректированного сигнала, параметры коррекции.

```
Recording started: 2025-4-26 0:25:33
End of the recording: 2025-4-26 0:25:33
Duration of the signal recorded (s) - 29
Input central frequency [Hz] - 957
Target central frequency [Hz] - 1408
High frequencies correction - 0
Target central frequency [Hz] - 0.938 x2
Input signal level - 4840
Target signal level - 2130
Signal Level Corr - 0.44
File - Speech Level-Max.wav
```

ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Мы рекомендуем пользователям время от времени заглядывать на наш сайт, в котором могут появляться полезные обновления SAT. Узнать больше о теоретических основах и программном коде разработки SAT: см. дополнительную информацию, размещаемую на этом сайте.