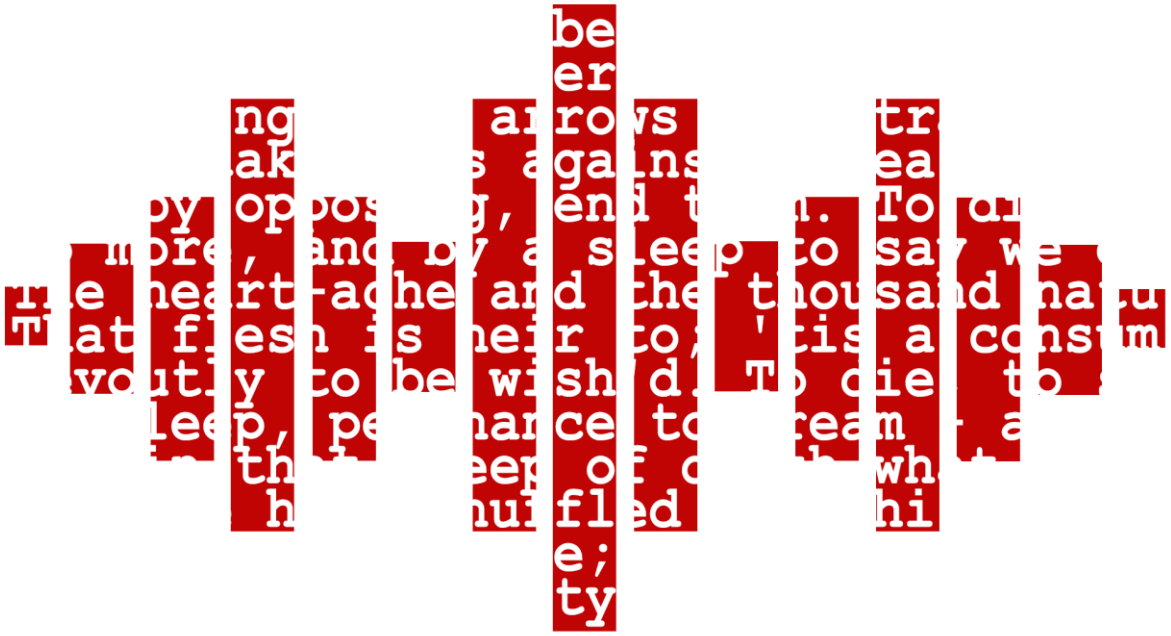
Speech Recognition System



stendhal.syndrome.studio@gmail.com

Contents

1. About…….3

2. Quick start……4

3. API overview……4

4. How to add the required language ……6

5. Contact us……6

1. About

Speech Recognition System is the plugin that provides high quality speech recognition offline.

Features:

- Does not require Internet connection;
- High quality and speed of speech recognition;
- Supports **17** different languages;
- Possibility to work with **Oculus Quest**;
- Multiplatformity;
- Easiness of integration.

Supported languages:

- English (download small model, download large model);
- Indian English (download small model, download large model);
- Chinese (download small model, download large model);
- Russian (download model);
- French (download small model, download large model);
- German (download small model, download large model);
- Spanish (download model);
- Portuguese (download model);
- Greek (download model);
- Turkish (download model);
- Vietnamese (download model);
- Italian (download model);
- Dutch (download small model, download large model);
- Catalan (download model);
- Arabic (download model);
- Farsi (download model);
- Filipino (download model).

Supported platforms:

- Windows 10, Windows 7 Service Pack 1;
- Android (armeabi-v7a or arm64-v8a).

Links:

Video instruction on plug-in usage

Demo. Windows build

Demo. Android build

Demo. Video

2. Quick start

- Import package **SpeechRecognitionSystem.unitypackage**.
- Make sure that the microphone is switched on and in working order;
- Enter settings of the microphone and adjust them: sensitivity is to be at zero;
- Load the scene **SampleScene** from folder **Demo**;
- Press button  to load the scene;
- Wait till there appear the words «Say something…» in the top right-hand corner of the screen. Say anything using the microphone (English is recognized by default). In real-time there will appear the Partial Result in the top right-hand corner and there will shortly appear the final result of speech recognition in the top left-hand corner.

3. API overview

3.1 AudioRecorder

Fields of class AudioRecorder:

- **Microphone Index** – the number (index) of the microphone that will be used during the work of the plugin.

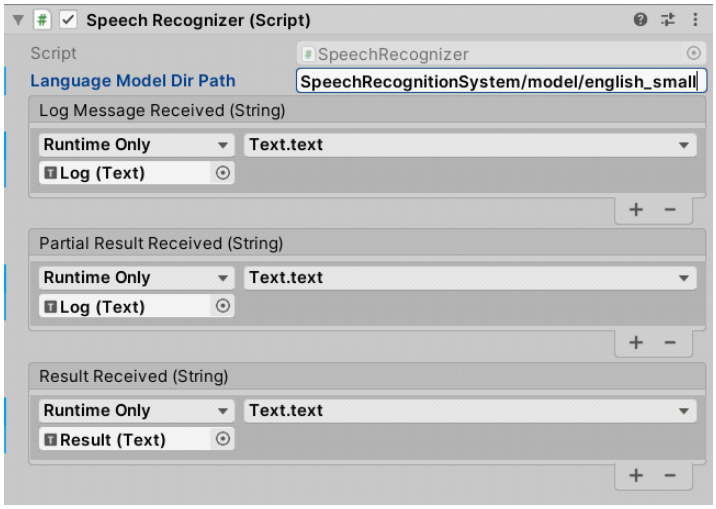
Class events of AudioRecorder:

- **MicReady (IMicrophone)** – event (**UnityEvent<IMicrophone>**) with results of initialization of the microphone (Unity Microphone). As the parameter the event transmits the object implementing the interface **IMicrophone.**

Methods of class AudioRecorder (implementation of IMicrophone):

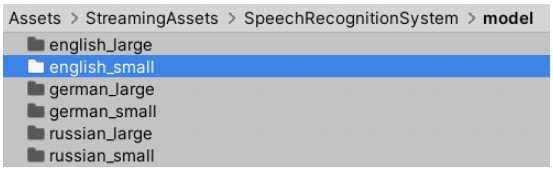
- **GetRecordPosition** – get the position in samples of the recording;
- **GetAudioClip** – the function returns null if the recording fails to start or the object of the AudioClip type if the recording has started successfully;
- **IsRecording** – query if a device is currently recording.

3.2 SpeechRecognizer



Fields of class SpeechRecognizer:

- **Language Model Dir Path** – the path to the directory with the files of the language model (in respect to Streaming Assets).



Methods of class SpeechRecognizer:

- **OnMicrophoneReady** – event handler **MicReady** of the component **AudioRecorder** of the object **AudioRecorder**. The method receives the object implementing the interface **IMicrophone** and saves the link to it for the further access to the microphone buffer;

Class events of SpeechRecognizer:

- **Log Message Received (String)** – event (**UnityEvent<bool>**) with log message.
- **Partial Result Received (String)** – event (**UnityEvent<String>**) with partial results of speech recognition;
- **Result Received (String)** – event (**UnityEvent<String>**) with results of speech recognition;

4. How to add the required language

Download the archive with the required language using the corresponding link provided below:

- English (download small model, download large model);
- Indian English (download small model, download large model);
- Chinese (download small model, download large model);
- Russian (download model);
- French (download small model, download large model);
- German (download small model, download large model);
- Spanish (download model);
- Portuguese (download model);
- Greek (download model);
- Turkish (download model);
- Vietnamese (download model);
- Italian (download model);
- Dutch (download small model, download large model);
- Catalan (download model);
- Arabic (download model);
- Farsi (download model);
- Filipino (download model).

For some languages there exist two types of language models: **small** one – for reasonable accuracy and high speed of speech recognition (for example, could be used for the work on mobile devices) and **large** one – for high accuracy and reasinable speed of speech recognition.

Unpack the contents of the downloaded archive to the directory **StreamingAssets/SpeechRecognitionSystem/model**. Specify the path to the language model in respect to **StreamingAssets** in the field **LanguageModelDirPath** of the component **SpeechRecognizer** of the object **SpeechRecognizer**. For example, «**SpeechRecognitionSystem/model/english_large**». Please pay attention that the line it not to end with «/» symbol.

5. Contact us

Do you meet issues while using this plugin?

Do you have suggestions how to improve API?

Feel free to contact us: **stendhal.syndrome.studio@gmail.com**