How to convert Speech to Text using C# with the help of Google Cloud Platform

Download: google-stt.zip

This example demonstrates how to implement the speech-to-text feature in c#, which is able to convert audio dat to text messages. The conversion is based on the powerful Google Cloud Speech API. The converted data can b an audio file, audio stream or real time human voice as well. Any audio supported by Ozeki VoIP SIP SDK i accepted. To understand this article, please read the following tutorial as well:
How to configure Google Cloud Platform to your Ozeki VoIP SDK projects

You can choose from all Google translation API supported languages.

An internet access is required. To try this example, you need to have Ozeki VoIP SIP SDK installed, and a reference t OzekiSDK.dll should be added to your Visual Studio project.

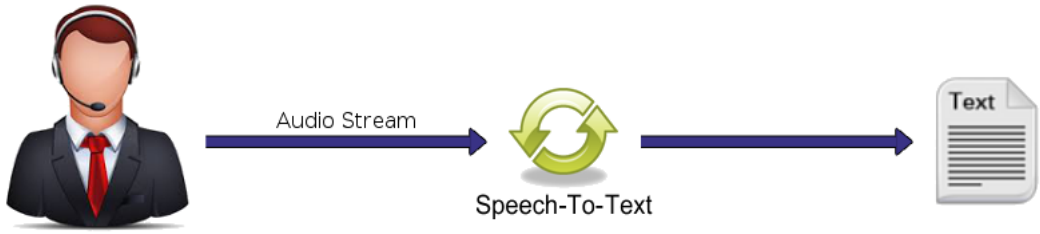

Figure 1 - Speech to Text conversion

What is Speech-to-Text used for?

A **speech-to-text** (STT) system converts normal speech from multiple languages into text. Users can set speech inputs an save them as text files, so later on the files can be read or analysed. You can use the text results for several purposes. Fo example you could store phone conversations in written forms. You can also store texts in an SQL database, forward them i e-mail or SMS or search keywords in them.

Speech-to-Text refers to the ability to listen to an audio stream and converting it to a text message. STT engines with differe languages, dialects and specialized vocabularies are available through the Google Cloud Speech API. Check if your require language is supported.

How to implement Google speech-to-text feature in your Ozeki VoIP SIP SDK project?

First you will need to register to the Google Cloud Platform, than you need to set the API access credentials on your operatin system and install the Google Cloud Speech SDK. After the installation is finished you will need to reboot your computer t test the example codes in Ozeki VoIP SIP SDK. **Here is a detailed tutorial on how you can set up and try you examples.**

The sample projects can be downloaded from here (google-stt.zip). Each project contains a basic example that combines th functionality of our SDK and the features provided by the Google Cloud Speech API, presented in a simple C# class GoogleSTT. The GoogleSTT class demonstrates how to implement Speech-to-Text functionality with OzekiSDK powered b the robust Google Cloud Speech API. A wide variety of languages can be given as parameter (e.g. an instance is created a line 25 to 27 in the C# example below). This instance can be attached to the call through the correct sender object (line 29 The instance in the current example can recognise United Kingdom English speech arriving through the microphone an converts it to text messages.

Microphone signals are converted to text in C# using the Google Cloud Speech API

'Program.cs'

```
1  using Ozeki.Media;
2  using System;
3
4  namespace Google_Speech_To_Text_V1
5  {
6      class Program
7      {
8          static MediaConnector connector;
9
```

```
        static Microphone microphone;

        static GoogleSTT googleSTT;

        public static void Main(string[] args)
        {
            Console.OutputEncoding = System.Text.Encoding.UTF8;

            connector = new MediaConnector();
            microphone = Microphone.GetDefaultDevice();

            var format = new WaveFormat(48000, 16, 1);

            microphone.ChangeFormat(format);

            googleSTT =
            new GoogleSTT(GoogleLanguage.English_United_Kingdom,
                                        format.AsVoIPMediaFormat());

            connector.Connect(microphone, googleSTT);

            microphone.Start();

            googleSTT.Start();

            Console.WriteLine("Speak !!");

            Console.ReadLine();

            Console.WriteLine("Disconnect");

            connector.Disconnect(microphone, googleSTT);

            Console.WriteLine("Google dispose");

            googleSTT.Dispose();
            googleSTT = null;

            Console.WriteLine("microphone dispose");

            microphone.Dispose();
            microphone = null;

            Console.WriteLine("connector dispose");

            connector.Dispose();
            connector = null;
        }
    }
}
```

'GoogleSTT.cs'

This 'GoogleSTT.cs' example class is capable to provide Speech-to-Text functionality through the Google Cloud Speech AP You can write classes similar to 'GoogleSTT.cs'.

From line 80 to 86 you can see the results of the speech-to-text conversation. The 'result.Alternatives' is a list of object containing every possible result and the confidence level of each result. When speech is converted by the Google Clou servers, the servers can understand speech as multiple texts and render a confidence level to each them from 0.0 to 1.0.

This example selects the text with the biggest confidence value and writes it on the console.

```
using Ozeki.Media;
using Google.Cloud.Speech.V1Beta1;

using System;
using System.Threading.Tasks;
using System.Threading;
using System.Linq;

namespace Google_Speech_To_Text_V1
{
    class GoogleSTT : AudioReceiver
    {
        SpeechClient speech;
        SpeechClient.StreamingRecognizeStream streamingCall;

        Task printResponses;

        AudioFormat _format;

        private string _languageCode;

        private GoogleLanguage _language;
        public GoogleLanguage Language
        {
            get { return _language; }
            set
            {
                _language = value;
                _languageCode = _language.GetCode();
            }
        }

        public GoogleSTT(string languageCode)
            : this(GoogleLanguageExt.GetGoogleLanguageFromCode(languageCode),
                    new AudioFormat())
        { }

        public GoogleSTT(GoogleLanguage languageCode, AudioFormat format)
```

```
        {
            Language = languageCode;

            SetReceiveFormats(format);

            _format = format;

            Init();
        }

        private void Init()
        {
            speech = SpeechClient.Create();

            streamingCall = speech.StreamingRecognize();

            streamingCall.WriteAsync(
               new StreamingRecognizeRequest()
               {
                   StreamingConfig = new StreamingRecognitionConfig()
                   {
                       Config = new RecognitionConfig()
                       {
                           Encoding =
                           RecognitionConfig.Types.AudioEncoding.Linear16,
                           SampleRate = _format.SampleRate,
                           LanguageCode = _languageCode,
                           MaxAlternatives = 5
                       },
                       InterimResults = true,
                   }
               });

            printResponses = Task.Run(async () =>
            {
                while (await streamingCall.ResponseStream.MoveNext(
                    default(CancellationToken)))
                {
                    foreach (var result in streamingCall.ResponseStream
                        .Current.Results)
                    {
                        if (result.IsFinal)
                        {
                            var top =
                            result.Alternatives.OrderBy(x => x.Confidence).First();

                            Console.WriteLine(top.Transcript);
                        }
                    }
                }
            });
        }

        object writeLock = new object();

        public bool IsRunning { get; private set; }

        public void Stop()
        {
            IsRunning = false;
        }

        public void Start()
        {
            IsRunning = true;
        }

        protected override void OnDataReceived(object sender, AudioData data)
        {
            if (!IsRunning) return;

            lock (writeLock)
            {
                var request = new StreamingRecognizeRequest();
                request.AudioContent = Google.Protobuf.ByteString
                            .CopyFrom(data.Data, 0, data.Data.Length);

                try
                {
                    streamingCall.WriteAsync(request).Wait();
                }
                catch (Exception e)
                {
                    streamingCall.WriteCompleteAsync();
                    Init();
                }
            }
        }

        protected override void Dispose(bool disposing)
        {
            Stop();

            if (printResponses != null)
            {
                printResponses = null;
            }

            if (streamingCall != null)
            {
                streamingCall = null;
            }

            if (speech != null)
```

```
143             {
144                 speech = null;
145             }
146
147             base.Dispose(disposing);
148         }
149     }
150 }
```

Related Pages

How to convert Text to Speech through Google

How to ring a SIP extension

How to send voice using microphone

How to make a voice call

Download Ozeki VoIP SDK

Ozeki VoIP SIP SDK - Training guides and examples

More information

- How to build a softphone voip sip client
- Register to SIP PBX
- Voip softphone development
- How to encrypt voip sip calls with sip encryption
- How to encrypt voip sip calls with rtp encryption
- How to ring a sip extension csharp example for sip invite
- How to make a sip voice call using csharp
- Voip multiple phone lines
- How to send stream of voice data into call using csharp microphone
- How to receive voice from SIP voice call using csharp speaker
- How to make conference voice call using voip sip
- How to play an mp3 file into a voice call using csharp
- How to convert text to speech and play that into a call using csharp
- How to use Microsoft Speech Platform 11 for TTS and STT
- How to record voip sip voice call
- How to accept incoming call using csharp
- How to reject incoming call using csharp
- How to read Headset buttons using Bluetooth
- How to implement auto answer using csharp
- How to recognize incoming voice using speech to text conversion
- Voip forward call
- Voip blind transfer
- Voip attended transfer
- Voip do not disturb
- Voip call hold
- SIP Message Waiting Indication
- Voip DTMF signaling
- How to work with sip and sdp in voip sip calls
- How to work with rtp in voip sip calls
- How to make voip video calls in csharp
- Voip video codec
- Google TextToSpeech

Video codecs | Install Google Cloud Speech