



- ▼ Introduction
- ▼ VoIP Technology
- ▼ Ozeki VoIP SDK
- **Developers Guide**
  - **Softphone Development**
    - **Basic softphone examples**
      - ▼ Build a softphone
      - ▼ Register to a SIP PBX
      - ▼ Ozeki VoIP SIP Softphone
      - ▼ SIP encryption
      - ▼ RTP encryption
      - ▼ Ring a SIP extension
      - ▼ Make a SIP voice call
      - ▼ Multiple phone lines
      - ▼ Voice via microphone
      - ▼ Receive call via speaker
      - ▼ Make conference call
      - ▼ Play mp3 into call
      - ▼ Use TextToSpeech
      - ▼ MS Speech Platform 11
      - ▼ Record voice call
      - ▼ Accept incoming call
      - ▼ Reject incoming call
      - ▼ Bluetooth Headset
      - ▼ Auto Answer
      - ▼ Voice Recognition
      - ▼ Forward incoming calls
      - ▼ Blind Transfer
      - ▼ Attended call transfer
      - ▼ Do not Disturb
      - ▼ Call holding
      - ▼ Show message waiting
      - ▼ Use DTMF signaling
      - ▼ Work with SIP and SDP
      - ▼ Work with RTP
      - ▼ Make video calls
      - ▼ Video codecs
    - **Google SpeechToText**
      - ▼ Install Google Cloud Speech
      - ▼ Google TextToSpeech
  - ▼ **Advanced examples**
  - ▼ **Sound quality**
  - ▼ **Direct calls among clients**
    - ▼ Manage NAT traversal
  - ▼ **Webphone Development**
  - ▼ **Voice Recording**
  - ▼ **IVR Development**
  - ▼ **PBX Development**
  - ▼ **Call Center Development**
  - ▼ **VoIP CRM Integration**
  - ▼ **Alert systems**
  - ▼ **IP Camera**
  - ▼ **Mobile phones and platforms**
  - ▼ **Billing**
  - ▼ **Tutorial**
  - ▼ **Appendix**
  - ▼ **FAQ**

## 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 data to text messages. The conversion is based on the powerful Google Cloud Speech API. The converted data can be an audio file, audio stream or real time human voice as well. Any audio supported by Ozeki VoIP SIP SDK is 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 OzekiSDK.dll](#) should be added to your Visual Studio project.

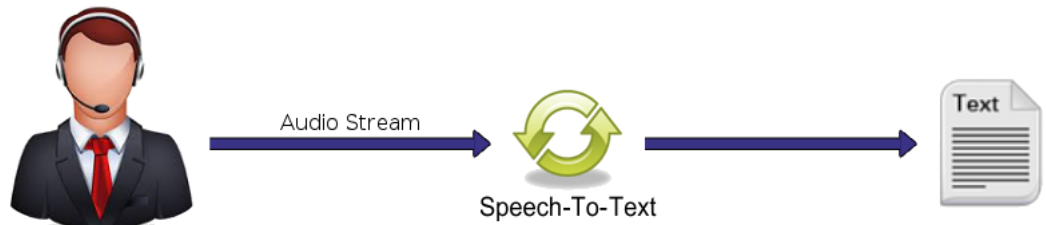


Figure 1 - Speech to Text conversion

## 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 and save them as text files, so later on the files can be read or analysed. You can use the text results for several purposes. For example, you could store phone conversations in written forms. You can also store texts in an SQL database, forward them via 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 different languages, dialects and specialized vocabularies are available through the Google Cloud Speech API. [Check if your required 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, then you need to set the API access credentials on your operating system and install the Google Cloud Speech SDK. After the installation is finished you will need to reboot your computer to test the example codes in Ozeki VoIP SIP SDK. [Here is a detailed tutorial on how you can set up and try your examples.](#)

The sample projects can be downloaded from [here \(google-stt.zip\)](#). Each project contains a basic example that combines the 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 by the robust Google Cloud Speech API. A wide variety of languages can be given as parameter (e.g. an instance is created at 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](#) and 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;
```

### RECOMMENDED FOR YOU:

**SMS Gateway software**  
Ozeki brings you outstanding **SMS Gateway** technology. Use our **SMS Server** products on Windows, Linux, or Android

**C# SMS API**

Developers can use our **C# SMS API** to send SMS from C#.Net. The C# SMS API comes with full source code

#### PHP SMS API

The ozeki **PHP SMS gateway** software can be used to send SMS from PHP and to receive SMS using PHP on your website

#### SMPP SMS Gateway

SMS service providers use our **SMPP gateway** solution, that offers a high performance **SMPP server** and **SMPP client** gateway with amazing routing capabilities

```

10 static Microphone microphone;
11
12 static GoogleSTT googleSTT;
13
14 public static void Main(string[] args)
15 {
16     Console.OutputEncoding = System.Text.Encoding.UTF8;
17
18     connector = new MediaConnector();
19     microphone = Microphone.GetDefaultDevice();
20
21     var format = new WaveFormat(48000, 16, 1);
22
23     microphone.ChangeFormat(format);
24
25     googleSTT =
26     new GoogleSTT(GoogleLanguage.English_United_Kingdom,
27                  format.AsVoIPMediaFormat());
28
29     connector.Connect(microphone, googleSTT);
30
31     microphone.Start();
32
33     googleSTT.Start();
34
35     Console.WriteLine("Speak !!");
36
37     Console.ReadLine();
38
39     Console.WriteLine("Disconnect");
40
41     connector.Disconnect(microphone, googleSTT);
42
43     Console.WriteLine("Google dispose");
44
45     googleSTT.Dispose();
46     googleSTT = null;
47
48     Console.WriteLine("microphone dispose");
49
50     microphone.Dispose();
51     microphone = null;
52
53     Console.WriteLine("connector dispose");
54
55     connector.Dispose();
56     connector = null;
57 }
58 }
59 }
```

#### 'GoogleSTT.cs'

This 'GoogleSTT.cs' example class is capable to provide Speech-to-Text functionality through the Google Cloud Speech API. 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 objects containing every possible result and the confidence level of each result. When speech is converted by the Google Cloud servers, the servers can understand speech as multiple texts and render a confidence level to each of them from 0.0 to 1.0.

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

```

1  using Ozeki.Media;
2  using Google.Cloud.Speech.V1Beta1;
3
4  using System;
5  using System.Threading.Tasks;
6  using System.Threading;
7  using System.Linq;
8
9  namespace Google_Speech_To_Text_V1
10 {
11     class GoogleSTT : AudioReceiver
12     {
13         SpeechClient speech;
14         SpeechClient.StreamingRecognizeStream streamingCall;
15
16         Task printResponses;
17
18         AudioFormat _format;
19
20         private string _languageCode;
21
22         private GoogleLanguage _language;
23         public GoogleLanguage Language
24         {
25             get { return _language; }
26             set
27             {
28                 _language = value;
29                 _languageCode = _language.GetCode();
30             }
31         }
32
33         public GoogleSTT(string languageCode)
34             : this(GoogleLanguageExt.GetGoogleLanguageFromCode(languageCode),
35                  new AudioFormat())
36         { }
37
38         public GoogleSTT(GoogleLanguage languageCode, AudioFormat format)
```

```

39     {
40         Language = languageCode;
41
42         SetReceiveFormats(format);
43
44         _format = format;
45
46         Init();
47     }
48
49     private void Init()
50     {
51         speech = SpeechClient.Create();
52
53         streamingCall = speech.StreamingRecognize();
54
55         streamingCall.WriteAsync(
56             new StreamingRecognizeRequest()
57             {
58                 StreamingConfig = new StreamingRecognitionConfig()
59                 {
60                     Config = new RecognitionConfig()
61                     {
62                         Encoding =
63                             RecognitionConfig.Types.AudioEncoding.Linear16,
64                         SampleRate = _format.SampleRate,
65                         LanguageCode = _languageCode,
66                         MaxAlternatives = 5
67                     },
68                     InterimResults = true,
69                 }
70             });
71
72         printResponses = Task.Run(async () =>
73         {
74             while (await streamingCall.ResponseStream.MoveNext(
75                 default(CancellationTokens)))
76             {
77                 foreach (var result in streamingCall.ResponseStream
78                     .Current.Results)
79                 {
80                     if (result.IsFinal)
81                     {
82                         var top =
83                             result.Alternatives.OrderBy(x => x.Confidence).First();
84
85                         Console.WriteLine(top.Transcript);
86                     }
87                 }
88             }
89         });
90     }
91
92     object writeLock = new object();
93
94     public bool IsRunning { get; private set; }
95
96     public void Stop()
97     {
98         IsRunning = false;
99     }
100
101     public void Start()
102     {
103         IsRunning = true;
104     }
105
106     protected override void OnDataReceived(object sender, AudioData data)
107     {
108         if (!IsRunning) return;
109
110         lock (writeLock)
111         {
112             var request = new StreamingRecognizeRequest();
113             request.AudioContent = Google.Protobuf.ByteString
114                 .CopyFrom(data.Data, 0, data.Data.Length);
115
116             try
117             {
118                 streamingCall.WriteAsync(request).Wait();
119             }
120             catch (Exception e)
121             {
122                 streamingCall.WriteCompleteAsync();
123                 Init();
124             }
125         }
126     }
127
128     protected override void Dispose(bool disposing)
129     {
130         Stop();
131
132         if (printResponses != null)
133         {
134             printResponses = null;
135         }
136
137         if (streamingCall != null)
138         {
139             streamingCall = null;
140         }
141
142         if (speech != null)

```

```
143     {  
144         speech = null;  
145     }  
146     base.Dispose(disposing);  
147 }  
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](#) ▶