

[AI & Machine Learning Products](https://cloud.google.com/products/machine-learning/) (https://cloud.google.com/products/machine-learning/)

[Cloud Speech-to-Text](https://cloud.google.com/speech-to-text/) (https://cloud.google.com/speech-to-text/)

[Documentation](https://cloud.google.com/speech-to-text/docs/) (https://cloud.google.com/speech-to-text/docs/) [Guides](#)

Selecting a transcription model

This page describes how to use a specific machine learning model for audio transcription requests to Speech-to-Text.

Transcription models

Speech-to-Text detects words in an audio clip by comparing input to one of many machine learning *models*. Each model has been trained by analyzing millions of examples—in this case, many, many audio recordings of people speaking.

Cloud Speech-to-Text has specialized models trained from audio from specific sources, for example phone calls or videos. Because of this training process, these specialized models provide better results when applied towards similar kinds of audio data.

For example, Cloud Speech-to-Text has a transcription model trained to recognize speech captured on a phone. When Cloud Speech-to-Text uses this model to transcribe phone audio, it produces significantly better results than other models.

The following table shows the transcriptions models available for use with Cloud Speech-to-Text.

Model name	Description
<code>command_and_search</code>	Best for short or single-word utterances like voice commands or voice search.
<code>phone_call</code>	Best for audio that originated from a phone call (typically recorded at an 8khz sampling rate).
<code>video</code>	Best for audio that originated from video or that includes more than one person talking. Ideally the audio is recorded at a 16khz or greater sampling rate. This is a premium model that costs more than the standard rate. See the pricing page (https://cloud.google.com/speech-to-text/pricing) for more details.
<code>default</code>	Best for audio that does not fit the other audio models, like long-form audio or

dictation. Ideally the audio is high-fidelity, recorded at a 16kHz or greater sampling rate.

Note: If you don't specify a model to use for speech recognition, Cloud Speech-to-Text attempts to select the model that best fits the settings in the [RecognitionConfig](#) of your request.

Selecting a model for audio transcription

To specify a specific model to use for audio transcription, you must set the `model` field to one of the allowed values—`video`, `phone_call`, `command_and_search`, or `default`—in the

[RecognitionConfig](#)

(<https://cloud.google.com/speech-to-text/docs/reference/rest/v1/RecognitionConfig>) parameters for the request. Cloud Speech-to-Text supports model selection for all speech recognition methods:

[speech:recognize](#) (<https://cloud.google.com/speech-to-text/docs/reference/rest/v1/speech/recognize>),

[speech:longrunningrecognize](#)

(<https://cloud.google.com/speech-to-text/docs/reference/rest/v1/speech/longrunningrecognize>), and

[Streaming](#)

(<https://cloud.google.com/speech-to-text/docs/reference/rpc/google.cloud.speech.v1#google.cloud.speech.v1.StreamingRecognizeRequest>)

.

PROTOCOL	NODE.JS	PYTHON	JAVA
<p>Refer to the speech:recognize (https://cloud.google.com/speech-to-text/docs/reference/rest/v1/speech/recognize) API endpoint for complete details.</p> <p>To perform synchronous speech recognition, make a <code>POST</code> request and provide the appropriate request body. The following shows an example of a <code>POST</code> request using <code>curl</code>. The example uses the access token for a service account set up for the project using the Google Cloud Cloud SDK (https://cloud.google.com/sdk). For instructions on installing the Cloud SDK, setting up a project with a service account, and obtaining an access token, see the quickstart (https://cloud.google.com/speech-to-text/docs/quickstart-protocol).</p> <pre>curl -s -H "Content-Type: application/json" \ -H "Authorization: Bearer \$(gcloud auth application-default print-access-token)" \ https://speech.googleapis.com/v1/speech:recognize \ --data "{ 'config': {</pre>			

```
'encoding': 'LINEAR16',
'sampleRateHertz': 16000,
'languageCode': 'en-US',
'model': 'video'
},
'audio': {
  'uri': 'gs://cloud-samples-tests/speech/Google_Gnome.wav'
}
}"
```

See the [RecognitionConfig](https://cloud.google.com/speech-to-text/docs/reference/rest/v1/RecognitionConfig)

(<https://cloud.google.com/speech-to-text/docs/reference/rest/v1/RecognitionConfig>) reference documentation for more information on configuring the request body.

If the request is successful, the server returns a **200 OK** HTTP status code and the response in JSON format:

```
{
  "results": [
    {
      "alternatives": [
        {
          "transcript": "OK Google stream stranger things from
Netflix to my TV okay stranger things from
Netflix playing on TV from the people that brought you
Google home comes the next evolution of the smart home
and it's just outside your window me Google know hi
how can I help okay no what's the weather like outside
the weather outside is sunny and 76 degrees he's right
okay no turn on the hose I'm holding sure okay no I'm can
I eat this lemon tree leaf yes what about this Daisy yes
but I wouldn't recommend it but I could eat it okay
Nomad milk to my shopping list I'm sorry that sounds like
an indoor request I keep doing that sorry you do keep
doing that okay no is this compost really we're all
compost if you think about it pretty much everything is
made up of organic matter and will return",
          "confidence": 0.9251011
        }
      ]
    }
  ]
}
```

Except as otherwise noted, the content of this page is licensed under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) (https://creativecommons.org/licenses/by/4.0/), and code samples are licensed under the [Apache 2.0 License](https://www.apache.org/licenses/LICENSE-2.0) (https://www.apache.org/licenses/LICENSE-2.0). For details, see our [Site Policies](https://developers.google.com/terms/site-policies) (https://developers.google.com/terms/site-policies). Java is a registered trademark of Oracle and/or its affiliates.