AI & Machine Learning Products　(https://cloud.google.com/products/machine-learning/)
Cloud Text-to-Speech　(https://cloud.google.com/text-to-speech/)
Documentation　(https://cloud.google.com/text-to-speech/docs/) Guides

Using device profiles for generated audio

This page describes how to select a device profile for audio created by Text-to-Speech.

You can optimize the synthetic speech
(https://cloud.google.com/text-to-speech/docs/basics#speech_synthesis) produced by Text-to-Speech
for playback on different types of hardware. For example, if your app runs primarily on smaller,
'wearable' types of devices, you can create synthetic speech from Cloud Text-to-Speech API that
is optimized specifically for smaller speakers.

You can also apply multiple device profiles to the same synthetic speech. The Cloud Text-to-
Speech API applies device profiles to the audio in the order provided in the request to the
`text:synthesize` (https://cloud.google.com/text-to-speech/docs/reference/rest/v1/text/synthesize)
endpoint. Avoid specifying the same profile more than once, as you can have undesirable
results by applying the same profile multiple times.

To hear the difference between audio generated from different profiles, compare the two clips
below.

> ▶ 0:00 / 0:05 ━━━━━━━━━━━ 🔊 ⋮
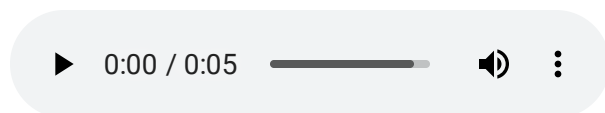
Example 1. Audio generated with `handset-class-device` *profile*

> ▶ 0:00 / 0:05 ━━━━━━━━━━━ 🔊 ⋮
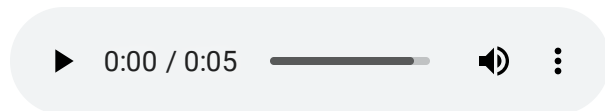
Example 2. Audio generated with `telephony-class-application` *profile*

Note: Each audio profile has been optimized for a specific device by adjusting a range of audio effects.
However, the make and model of the device used to tune the profile may not match users' playback devices
exactly. You may need to experiment with different profiles to find the best sound output for your application.

Available audio profiles

The following table gives the IDs and examples of the device profiles available for use by the Cloud Text-to-Speech API.

Audio profile ID	Optimized for
`wearable-class-device`	Smart watches and other wearables, like Apple Watch, Wear OS watch
`handset-class-device`	Smartphones, like Google Pixel, Samsung Galaxy, Apple iPhone
`headphone-class-device`	Earbuds or headphones for audio playback, like Sennheiser headphones
`small-bluetooth-speaker-class-device`	Small home speakers, like Google Home Mini
`medium-bluetooth-speaker-class-device`	Smart home speakers, like Google Home
`large-home-entertainment-class-device`	Home entertainment systems or smart TVs, like Google Home Max, LG TV
`large-automotive-class-device`	Car speakers
`telephony-class-application`	Interactive Voice Response (IVR) systems

Specify an audio profile to use

To specify an audio profile to use, set the `effectsProfileId` field for the speech synthesis request.

PROTOCOL	C#	GO		MORE ▾

To generate an audio file, make a `POST` request and provide the appropriate request body. The following shows an example of a `POST` request using `curl`. The example uses the access token for a service account set up for the project using the Google Cloud Platform 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/text-to-speech/docs/quickstart).

The following example shows how to send a request to the `text:synthesize` (https://cloud.google.com/text-to-speech/docs/reference/rest/v1/text/synthesize) endpoint.

```
curl \
  -H "Authorization: Bearer "$(gcloud auth print-access-token) \
  -H "Content-Type: application/json; charset=utf-8" \
  --data "{
    'input':{
      'text':'This is a sentence that helps test how audio profiles can change the
    },
    'voice':{
      'languageCode':'en-us',
    },
    'audioConfig':{
      'audioEncoding':'LINEAR16',
      'effectsProfileId': ['telephony-class-application']
    }
  }" "https://texttospeech.googleapis.com/v1beta1/text:synthesize" > audio-profile.
```

If the request is successful, the Cloud Text-to-Speech API returns the synthesized audio as base64-encoded data contained in the JSON output. The JSON output in the `audio-profiles.txt` file looks like the following:

```
{
  "audioContent": "//NExAASCCIIAAhEAGAAEMW4kAYPnwwIKw/BBTpwTvB+IAxIfghUfW.."
}
```

To decode the results from the Cloud Text-to-Speech API as an MP3 audio file, run the following command from the same directory as the `audio-profiles.txt` file.

```
sed 's|audioContent| |' < audio-profile.txt > tmp-output.txt && \
tr -d '\n ":{}' < tmp-output.txt > tmp-output-2.txt && \
base64 tmp-output-2.txt --decode > audio-profile.wav && \
rm tmp-output*.txt
```