

Documentation for Text to Speech API for IT IS 6177

API name:- Microsoft Azure Text to Speech API

Project is built on the Microsoft Azure Text to Speech API as base.

Requirements

- [Node.js](#)
- [Postman](#)

Setup

- Create a free [Azure](#) account.
- Create a speech resource and get subscription key and endpoint

Installation

- git clone <https://github.com/akashnevtia/Final-Project>
- cd Final-Project

User needs to create a .env file and paste following data

- API_URL = <https://eastus.tts.speech.microsoft.com/cognitiveservices/v1>
- API_AUTH_URL = <https://eastus.api.cognitive.microsoft.com/sts/v1.0/issuetoken>
- API_KEY = <Your API Key>
- SERVICE_NAME_NEURAL = (en-US, JessaNeural)
- VOICE_URL = <https://eastus.tts.speech.microsoft.com/cognitiveservices/voices/list>
- VOICE_HOST = eastus.tts.speech.microsoft.com

Install dependencies

- npm install

Start the application

- node app.js

Testing and Using application

It can be tested on Postman at <http://localhost:3000/>

Usage: -

There are two endpoints for this API.

- /voices
- /generate

/voices

This is an GET API the gives the all the available list of voices.

This is an sample output: -

```
{
  "Name": "Microsoft Server Speech Text to Speech Voice (af-ZA, AdriNeural)",
  "DisplayName": "Adri",
  "LocalName": "Adri",
  "ShortName": "af-ZA-AdriNeural",
  "Gender": "Female",
  "Locale": "af-ZA",
  "LocaleName": "Afrikaans (South Africa)",
  "SampleRateHertz": "24000",
  "VoiceType": "Neural",
  "Age": "18-24",
  "Gender": "Female",
  "Locale": "af-ZA",
  "LocaleName": "Afrikaans (South Africa)",
  "SampleRateHertz": "24000",
  "VoiceType": "Neural"
}
```

The parameter **ShortName** is very important. This is supposed to be passed in the body for user defined voice.

/generate

This is an POST API to convert the text to speech. We need to include body in the request.

Here is an example of how the body would look: -

```
{
  "TEXT": "type what you wanna listen",
  "VOICE": "en-US-AnaNeural",
  "SPEAKING_SPEED": 1,
  "PITCH": 0
}
```

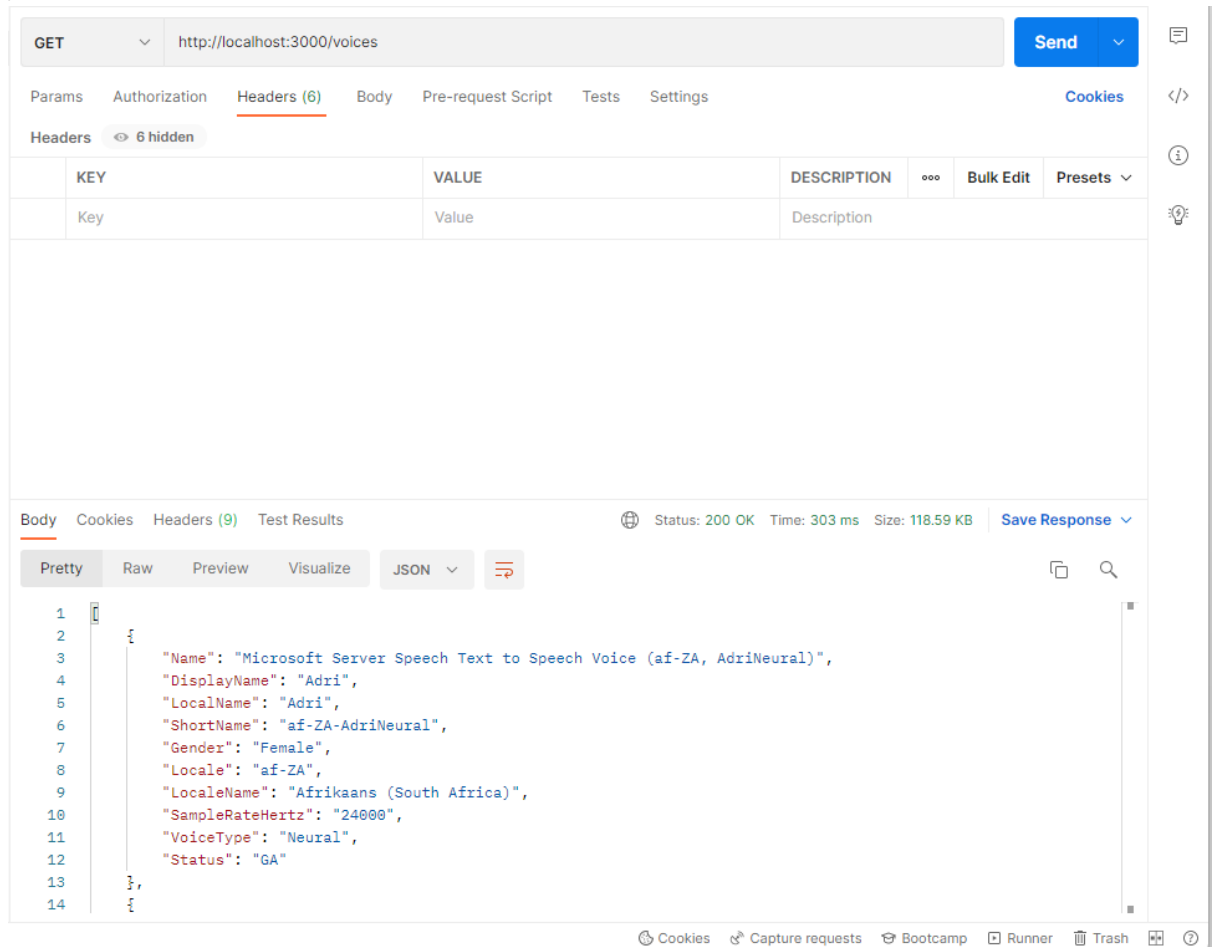
Keep in mind that the parameters are case sensitive. The left-hand side “key parameter” are case sensitive and needs to be entered as shown.

Rules for parameters

- “TEXT” this can be anything the needs to be converted to speech.
- “VOICE” this should be exactly as shown in /voices API. The ShortName parameter contains the value for this attribute
- “SPEAKING_SPEED” this defines the speed of speaker. It ranges from 0 to 3
- “PITCH” this defines the pitch of speaker. It ranges from (-50 to +50)
- If any values are not entered default values are assigned

Examples using Postman

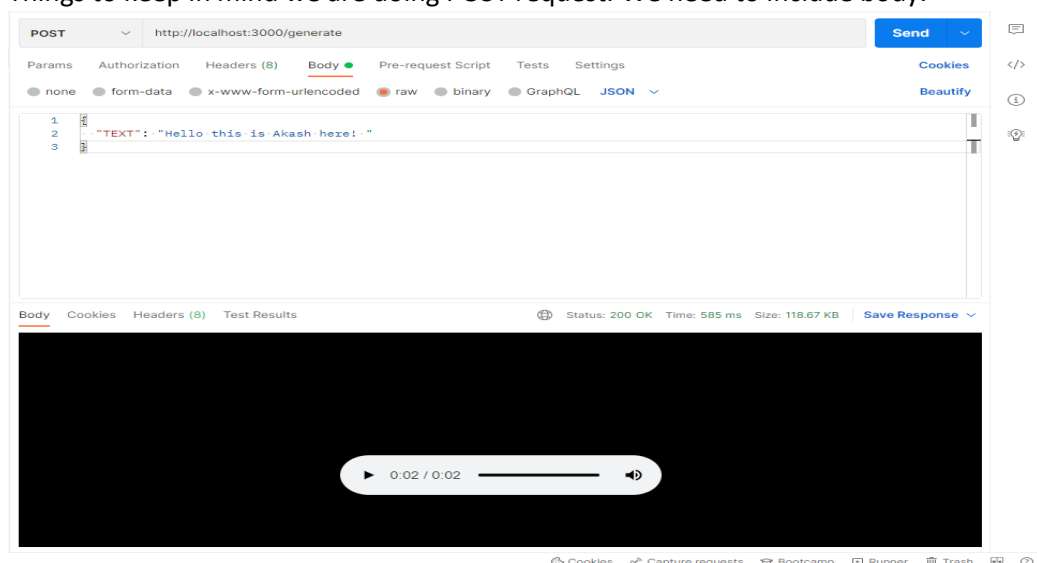
- /voices



We don't need to enter any details no Header just send the GET request.

/generate example

- First, we go with only "TEXT" as input
- Things to keep in mind we are doing POST request. We need to include body.



Next are a few examples

The screenshot displays a REST client interface with the following details:

- Method:** POST
- URL:** http://localhost:3000/generate
- Body Type:** JSON
- Body Content:**

```
1 {  
2   "TEXT": "Hello this is Akash here! ",  
3   "VOICE": "ur-IN-GulNeural"  
4 }
```
- Status:** 200 OK
- Time:** 584 ms
- Size:** 127.36 KB
- Response Type:** Video (indicated by the video player interface)

POST

http://localhost:3000/generate

Send

Params

Authorization

Headers (8)

Body

Pre-request Script

Tests

Settings

none

form-data

x-www-form-urlencoded

raw

binary

GraphQL

JSON

1

2

3

4

5

TEXT

TEXT

VOICE

SPEAKING_SPEED

1

2

3

4

5

1

2

3

4

5

Body

Cookies

Headers (8)

Test Results

Status: 200 OK

Time: 1355 ms

Size: 254.41 KB

Save Response

0:05 / 0:05

Cookies

Capture requests

Bootcamp

Runner

Trash

POST

http://localhost:3000/generate

Send

Params

Authorization

Headers (8)

Body

Pre-request Script

Tests

Settings

none

form-data

x-www-form-urlencoded

raw

binary

GraphQL

JSON

1

2

3

4

5

6

TEXT

TEXT

VOICE

SPEAKING_SPEED

PITCH

1

2

3

4

5

6

1

2

3

4

5

6

Body

Cookies

Headers (8)

Test Results

Status: 200 OK

Time: 457 ms

Size: 127.27 KB

Save Response

0:02 / 0:02

Cookies

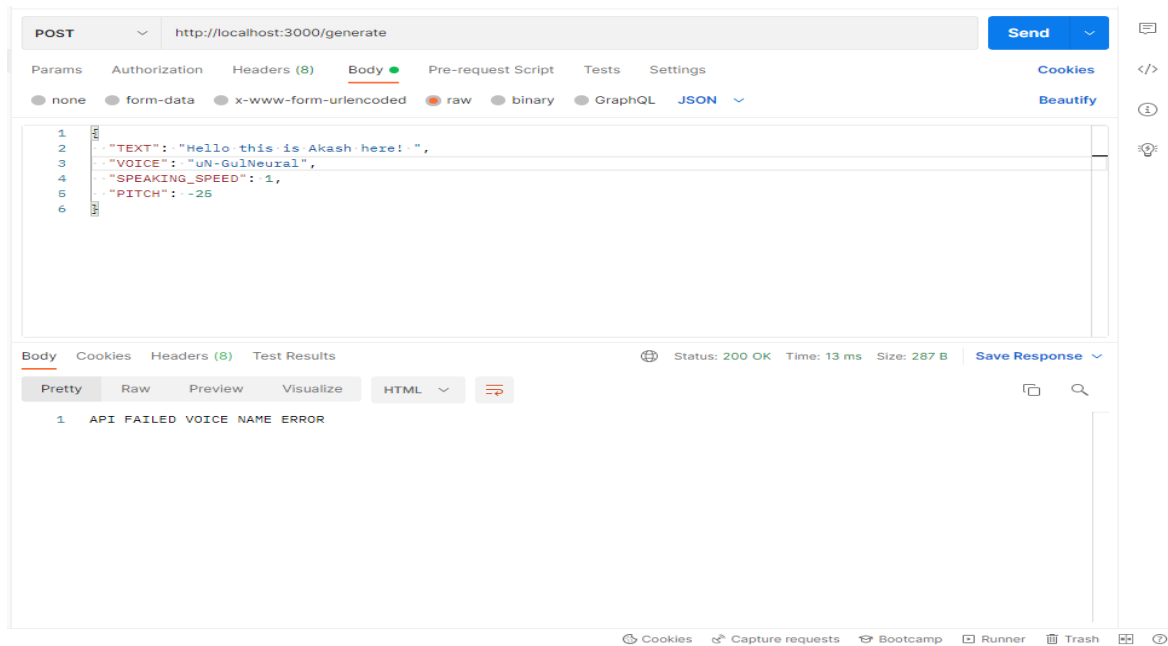
Capture requests

Bootcamp

Runner

Trash

This were the basic examples. Below is an example of voice error if wrong input is given.



Swagger UI for the same can be accessed at <http://localhost:3000/docs>

References

[Microsoft Azure API](#)

[Swagger](#)

[Node.js](#)

[Postman](#)