

Step 1: Install necessary packages

```
!pip install TTS
```

```
!pip install IPython
```

```
!pip install google.colab
```

```
!pip install pydub # To handle MP3 to WAV conversion
```

Step 2: Import necessary libraries

```
from google.colab import files
```

```
import IPython.display as ipd
```

```
from TTS.api import TTS
```

```
from pydub import AudioSegment # Library to convert mp3 to wav
```

```
import os
```

Step 3: Open the interface to upload the file (your voice model)

```
uploaded = files.upload()
```

Get the path of the uploaded file dynamically

```
for filename in uploaded.keys():
```

```
    wav_path = "/content/" + filename # Indentation applied here
```

```
    # Check if the file is MP3 and convert it to WAV
```

```
    if wav_path.endswith('.mp3'):
```

```
        audio = AudioSegment.from_mp3(wav_path)
```

```
        wav_output_path = wav_path.replace(".mp3", ".wav")
```

```
        audio.export(wav_output_path, format="wav")
```

```
        print(f"✅ File '{wav_path}' uploaded and converted to WAV as '{wav_output_path}'")
```

```
    else:
```

```
        print("❌ The uploaded file is not an MP3.")
```

Step 4: Load the XTTS model

```
tts = TTS("tts_models/multilingual/multi-dataset/xtts_v2").to("cpu")
```

Or

Load a TTS model that is designed for American English

```
tts = TTS("tts_models/en/ljspeech/tacotron2-DCA").to("cpu")
```

Step 5: Input text to generate audio

Input field for text to generate audio

```
text_input = input("Enter the text you want to clone: ")
```

```
# Step 6: Function to generate the audio with the cloned voice and play it
```

```
def generate_audio(text):
```

```
    # Generate the audio with the cloned voice
```

```
    output_path = "/content/voz_clonada.wav" # Path for the generated audio file
```

```
    tts.tts_to_file(
```

```
        text=text, # Text to clone
```

```
        file_path=output_path, # Output file path
```

```
        speaker_wav=wav_output_path, # Original voice file for cloning (WAV format)
```

```
        # language="auto" # Automatically detect the language of the text
```

```
    )
```

```
# Confirm the audio generation
```

```
print(f"✅ Audio synthesized and saved as '{output_path}')
```

```
# Verify that the file exists
```

```
if os.path.exists(output_path):
```

```
    print(f"✅ The file '{output_path}' exists.")
```

```
else:
```

```
    print(f"❌ The file '{output_path}' was not generated.")
```

```
# Play the generated audio to test
```

```
try:
```

```
    audio_player = ipd.Audio(output_path) # Create an audio player
```

```
    display(audio_player) # Display the audio player in Colab
```

```
    print(f"✅ Audio is ready to be played: {output_path}")
```

```
except Exception as e:
```

```
    print(f"❌ Error playing audio: {str(e)}")
```

```
# Step 7: Generate the audio using the entered text
```

```
generate_audio(text_input)
```