

✓ Ejercicio con SpeechRecognition

Realizar por medio de un .wav un SpeechRecognition

```
# Instalacion global en colab
!pip install SpeechRecognition
```

```
🔄 Collecting SpeechRecognition
  Downloading speechrecognition-3.14.3-py3-none-any.whl.metadata (30 kB)
  Requirement already satisfied: typing-extensions in /usr/local/lib/python3.11/dist-packages (from SpeechRecognition) (4.13.2)
  Downloading speechrecognition-3.14.3-py3-none-any.whl (32.9 MB)
    _____ 32.9/32.9 MB 24.4 MB/s eta 0:00:00
  Installing collected packages: SpeechRecognition
  Successfully installed SpeechRecognition-3.14.3
```

```
# Importando librerias
import speech_recognition as sr
```

```
# Creamos la instancia
r = sr.Recognizer()
```

✓ Trabajando con el fichero de audio

Para este caso usaremos un .wav

[De un dialogo de un capitulo de un show animado](#)

```
test_audio_file = sr.AudioFile('test_audio.wav')

with test_audio_file as source:
    test_audio = r.record(source)

type(test_audio)
```

**speech_recognition.audio.AudioData**

```
def __init__(frame_data, sample_rate, sample_width)
```

/usr/local/lib/python3.11/dist-packages/speech_recognition/audio.py

Creates a new ``AudioData`` instance, which represents mono audio data.

The raw audio data is specified by ``frame_data``, which is a sequence of bytes representing a

The width of each sample, in bytes, is specified by ``sample_width``. Each group of ``sample_w

```
# Usamos el api de google para la convesion
r.recognize_google(test_audio, language='es-ES')
```



```
'es el martillo Es real tómenme una fotocopia Oh No hermano'
```

Conclusion

Con este ejercicio se demostró que convertir audio a texto no es tan complicado como parece. Usando la librería `speech_recognition` y un simple archivo `.wav`, se pudo extraer lo que se decía en el audio con solo unas líneas de código. La herramienta de Google hizo el trabajo pesado, y la transcripción salió bastante bien. Esto puede servir para cosas como analizar audios, hacer subtítulos automáticos o simplemente experimentar con reconocimiento de voz. Lo bueno es que no se necesita saber tanto para que funcione, y eso lo vuelve bastante útil.

