

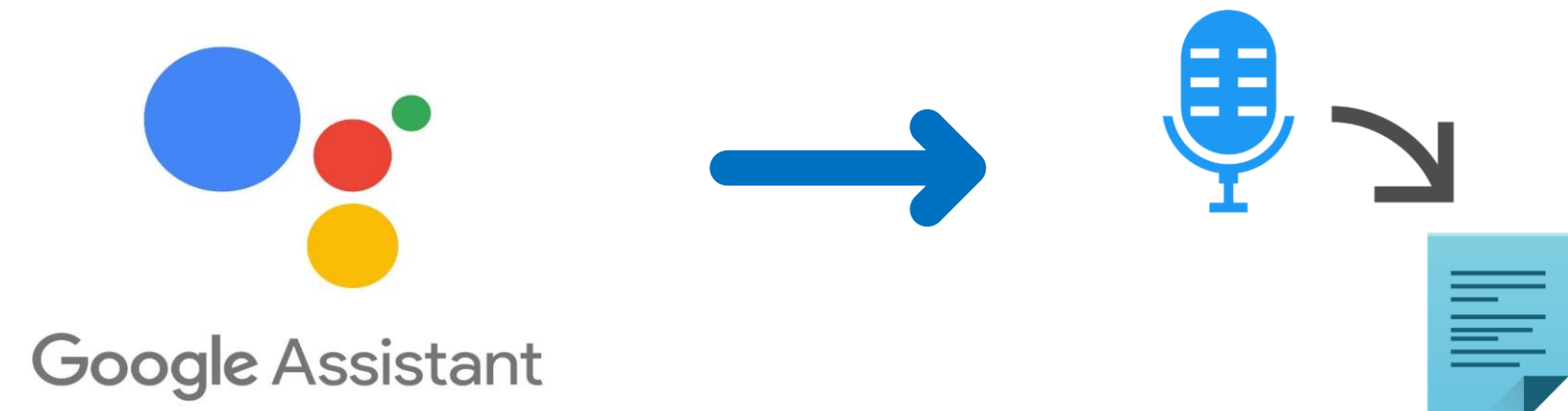
Abstract

Speech recognition is a tool that can convert speech to text. This will save people time and benefit people who do not know how to write in English or other languages. They should be able to speak, and the program will understand them. It can help with translations; it can be used to understand different accents. In addition, people with disabilities that make it difficult for them to type can use this program to communicate with the world.

Introduction

Most people use Google, and they have Google Assistant installed on their devices, such as smartphones. It allows us to talk with the robot and it will assist us by answering questions and controlling devices on the phone.

Google can convert the user's speech to text using Natural Language Processing and Neural Networks.



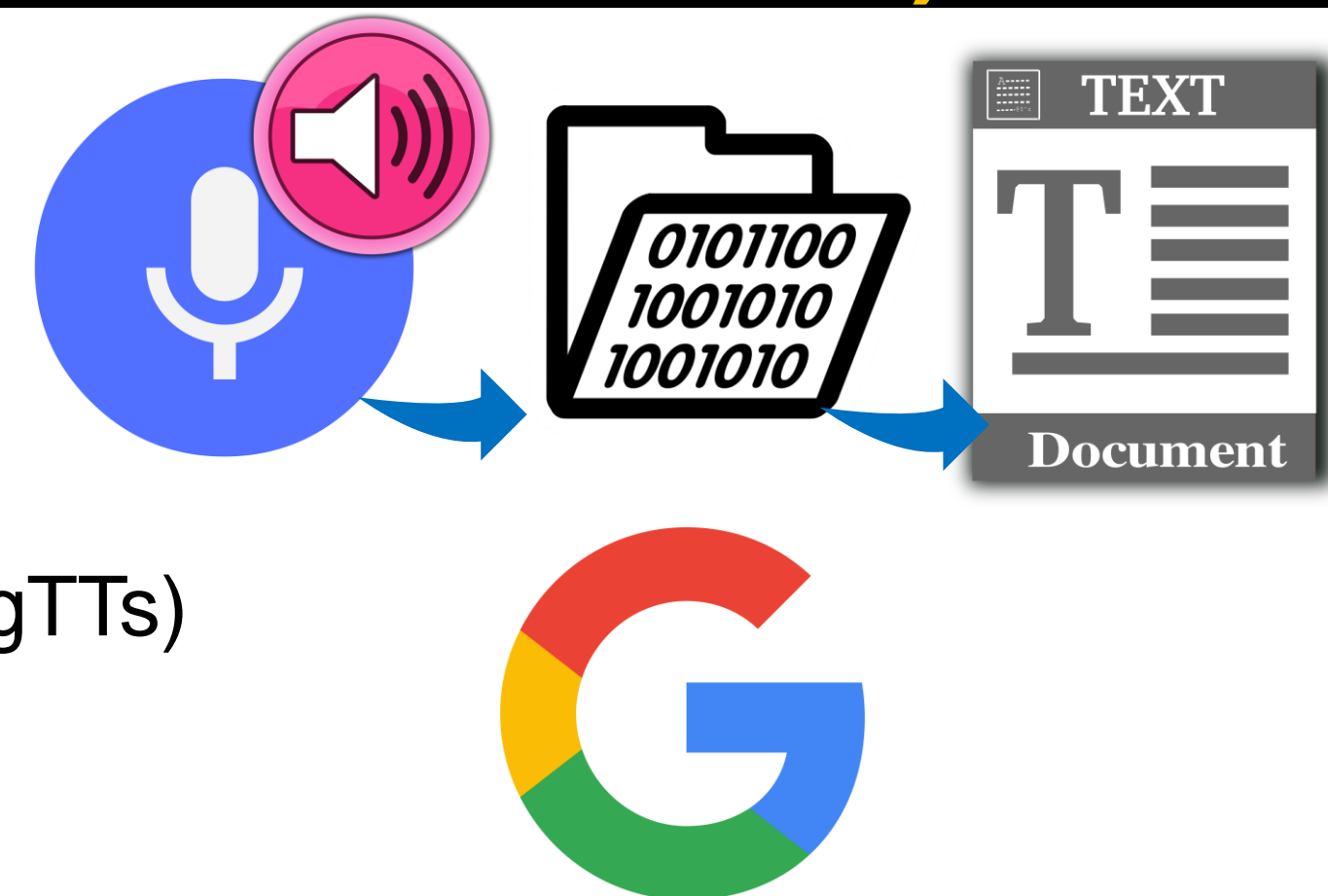
This project makes use of speech recognition in Python libraries from Google, such as Google Speech-to-Text. The program first installs the dependencies that are required to run the program. The program then asks the user for their input. The user has the option of either exit the program, having the text read back to them, or responding to the conversation. If the user decides to have the text read back to them, the program will use the Google Text-to-Speech library to convert the text to speech. The purpose of this project is to understand how to implement speech recognition in python and apply the concepts learned from programming II

Research Question(s)

- How to teach beginners to understand the basic concept of speech recognition?
- How does the Google speech recognition library work in python?
- Which concepts of the programming II class at KSU apply to this project?

Materials (Tools and Libraries)

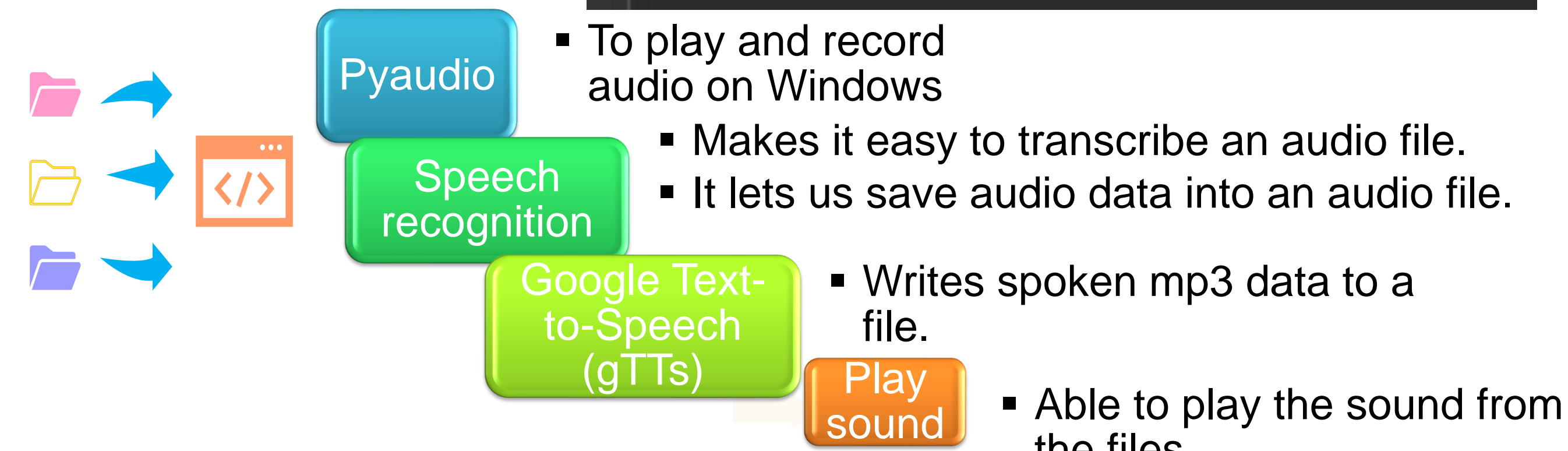
- Windows 10
- Microphone and speaker
- Install: Pyaudio
- Import:
 - Google Text-to-Speech (gTTS)
 - Speech recognition
 - Play sound



Methods

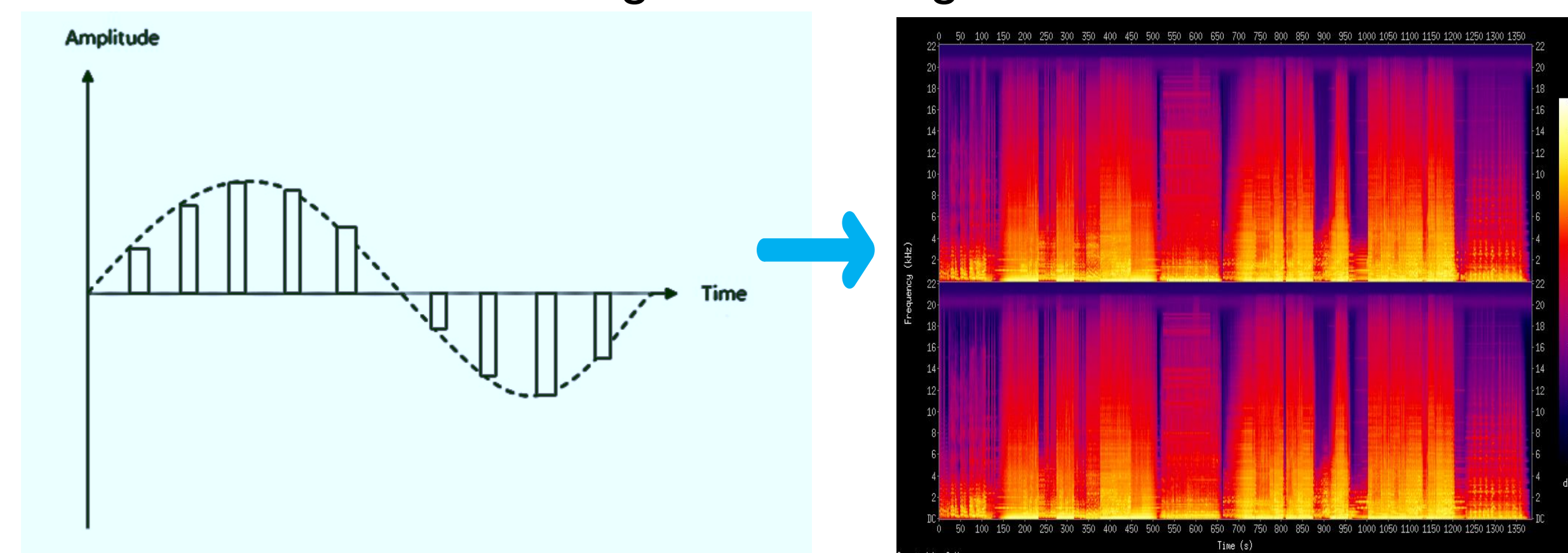
1. Importing all libraries

```
import speech_recognition as sr
from gtts import gTTS
import winsound
from playsound import playsound
from datetime import datetime
```



2. Setting the microphone and taking in the voice input

1. Convert the analog sound to digital

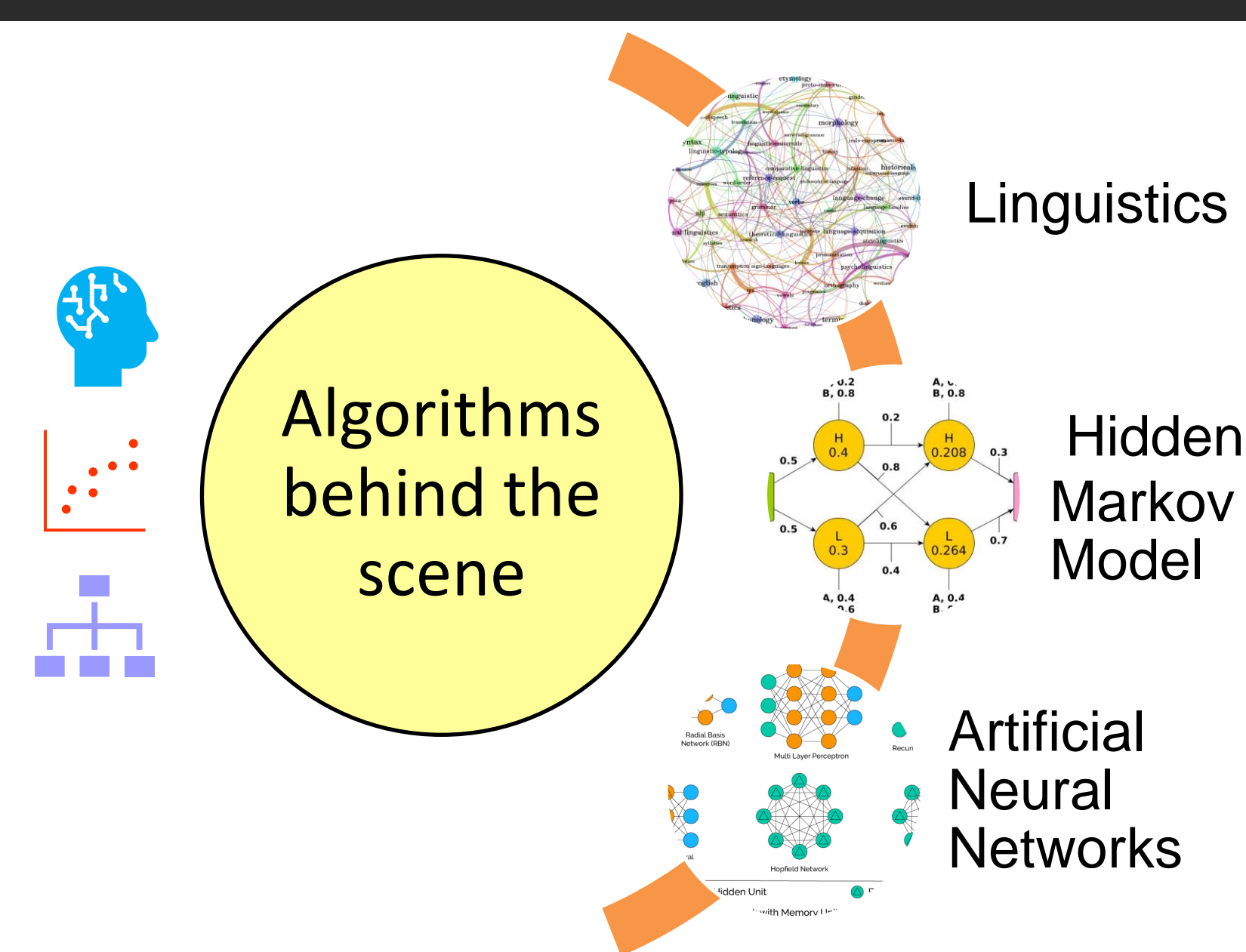


2. Fast Fourier Transform converts the graph above into a spectrogram

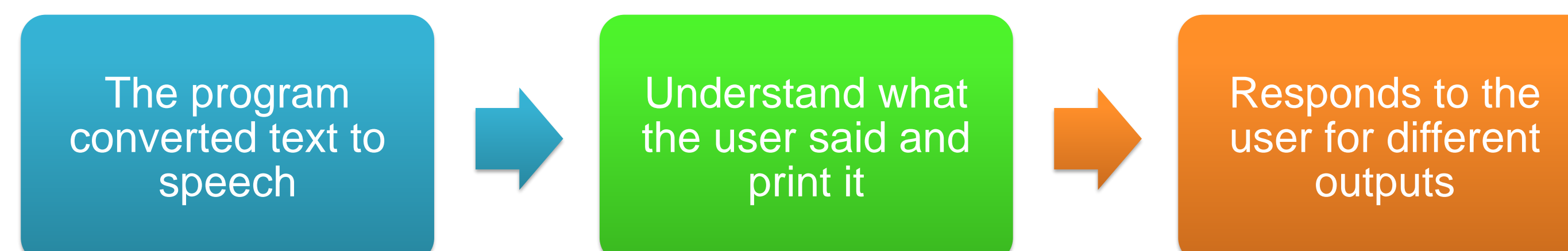
3. Converting the audio to text

```
audio = your_voice.listen(source) # Get the voice
text = your_voice.recognize_google(audio, language='en')

voice_text = gTTS(text)
print(text)
```



Results



Conclusions

Speech recognition is complex and hard to understand for the beginner. However, there is a library for a beginner to use for translating audio to text, which is known as Google Text-to-Speech (gTTS). This library uses a hidden Markov model (HMM) to recognize spoken words. HMMs are a type of statistical model that is often used in speech recognition and other tasks involving sequential data. It is trained on a large corpus of spoken English, which means that it can recognize a wide variety of accents and dialects. In addition to recognition, the library can also be used to generate text-to-speech output. This project is practicing the use of algorithms, handling exceptions, natural language processing, and neural networks.

Acknowledgments

I would like to thank my advisor, Professor Nick Murphy who has supported and guided me throughout this project. I am also grateful for my friend who helped me record the video. As well as the members of the Autonomous Underwater Vehicle team for letting me participate in the team and gaining knowledge from them.

Contact Information

Rachnicha Rojhanarittikorn: rrojghan@students.kennesaw.edu
<https://www.linkedin.com/in/rachnicha-rojhanarittikorn>

References

MathWorks. *Hidden Markov Models (HMM)*. 2022. <https://www.mathworks.com/help/stats/hidden-markov-models-hmm.html>.

S, Ravikiran A. *A Guide to Speech Recognition in Python: Everything You Should Know*. 27 September 2022. <https://www.simplilearn.com/tutorials/python-tutorial/speech-recognition-in-python>.

Vidyalaya, CS. *Signal Conversion: Analog to Digital Conversion*. 2022. <https://csvidyalaya.com/signal-conversion-analog-to-digital-conversion/>.

Williams, Carlos. *How to know if an audio file is really lossless using its spectrogram?* 16 February 2015. <https://dsp.stackexchange.com/questions/20563/how-to-know-if-a-audio-file-is-real-lossless-using-its-spectrogram>.