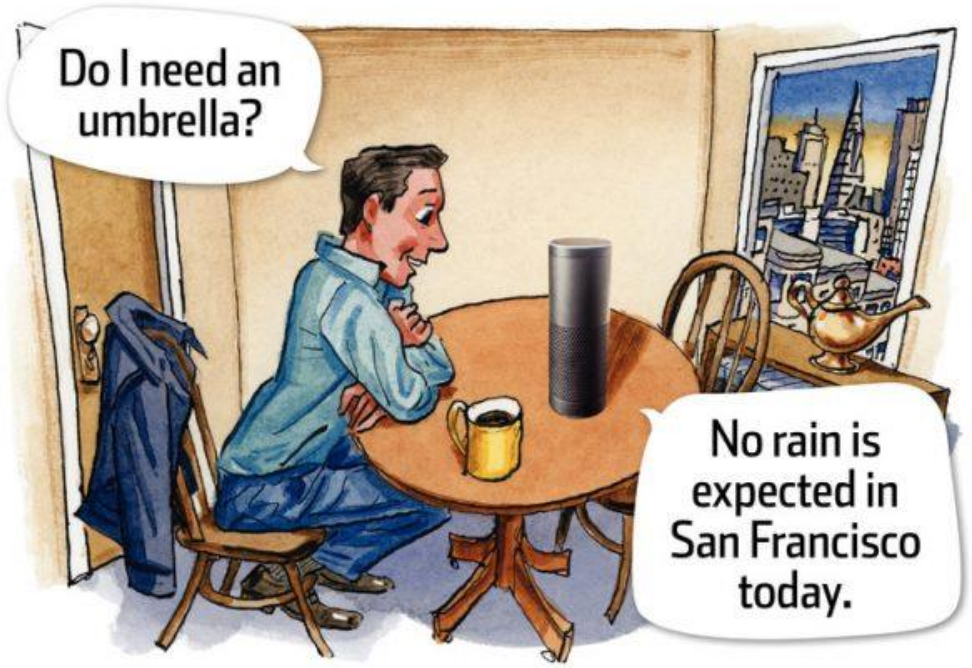
Speech Recognition

Do you know how Alexa or other personal assistants work?



Well, don’t you worry! You’ll learn that in this lesson.

Key Learning Outcomes

At the end of the lesson, you will be able to:

1. Understand how speech recognition works.
2. Use AI blocks in PictoBlox to convert speech into text.
3. Make your own virtual assistant in PictoBlox which recognizes your that can recognize your command play the requested song.

How Speech Recognition Works

How Do Humans Learn a Language?

From the time we are born, we hear words and sounds around us. Even before we can speak, we hear some words that we start responding to words like Mama, Dada, Yes, No.

Our brain tries to find patterns to differentiate various sounds and words and categorize them. It may seem as though humans are pre-programmed to listen and understand but it is not so. We have been trained to develop this ability.

Speech recognition technology has been developed on the same lines. Computers are also trained in the same way.

Speech Recognition

> *Speech recognition is the ability of a machine to identify words and phrases in spoken language and convert them to a machine-readable format.*

How Speech Recognition Works

Speech recognition is very complex and a lot of mathematical equations are involved. Let's break it down into simple steps:

1. First, the machine records the audio file.
2. Then, it breaks down the audio to extract consonants and vowels (the building blocks of a text). After this process, we get a list of consonants and vowels.
3. Using the word database of the language, the machine tries to identify words from the list and then make sentences thus converting the speech into text.

How Alexa Works

Alexa, Amazon's virtual assistant AI technology, uses *natural language processing*, a procedure of converting speech into sounds, words, and ideas.

Here's how she works:

1. Alexa first records your speech. Then, this recording is sent to Amazon's servers to be analyzed more efficiently.
2. Amazon breaks down the recording into individual sounds. It then consults a database containing various words' pronunciations to find which words most closely correspond to the combination of individual sounds.
3. It then identifies keywords to make sense of the tasks and carry out corresponding functions. E.g. if Alexa notices words like "weather" or "temperature", it will open the weather app.
4. Amazon's servers send the information back to your device. If Alexa needs to say anything back to you, it will go through the same process described above, but in reverse order.

Speech-to-Text Blocks in PictoBlox

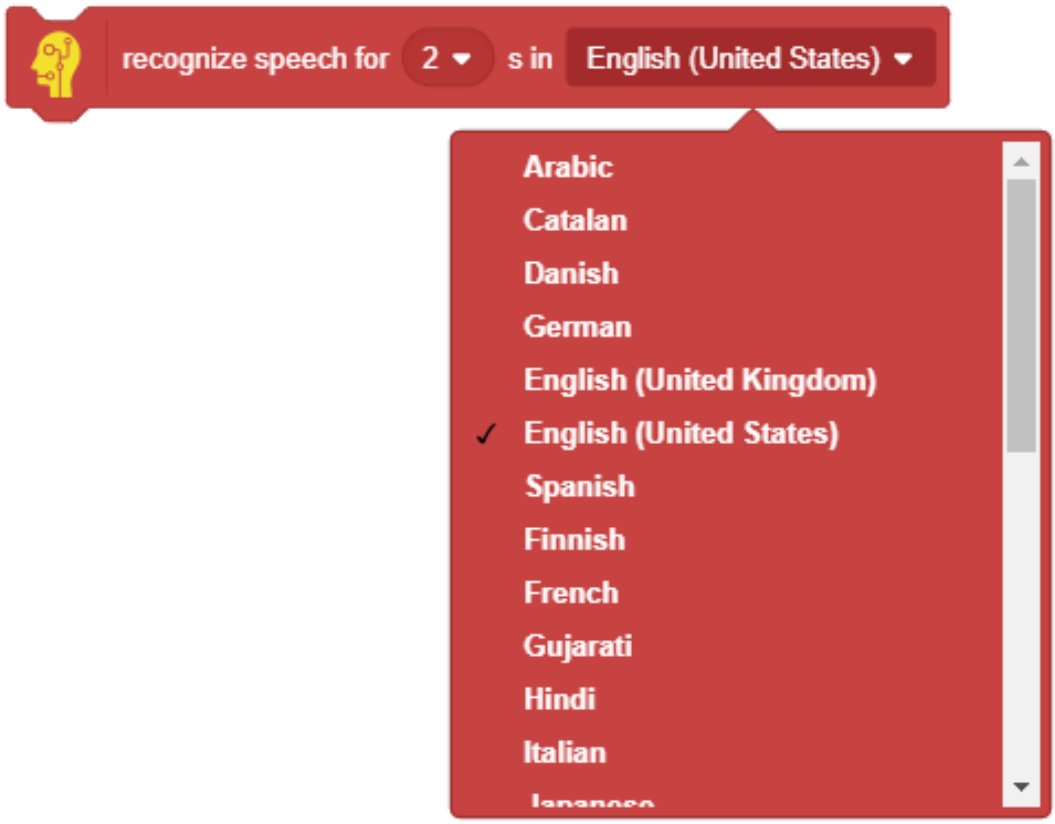
Speech-to-Text Blocks in PictoBlox

The *Artificial Intelligence* extension in PictoBlox has blocks dedicated to speech recognition. Let's first add the extension in our project:

1. Create a new project in PictoBlox.
2. Select evive as your board from the Board tab in the menu bar.
3. Next, click on the *Add Extension* button and add Artificial Intelligence extension.

Speech Recognition Block

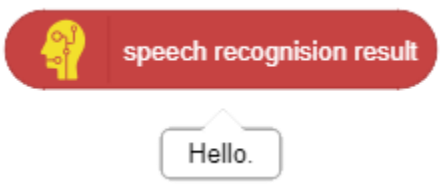
To execute speech recognition, we have the recognize speech for () s in ().



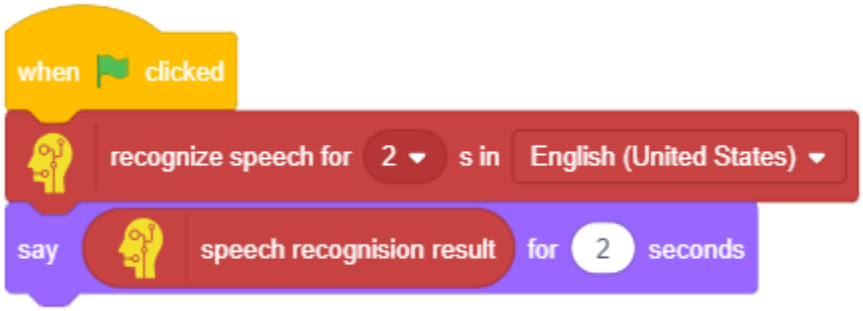
When the block is executed, the recognition window will open and you will get a specified time during which PictoBlox will record whatever you say. Once recorded, the speech will be converted to the text of the language you spoke in and saved locally.

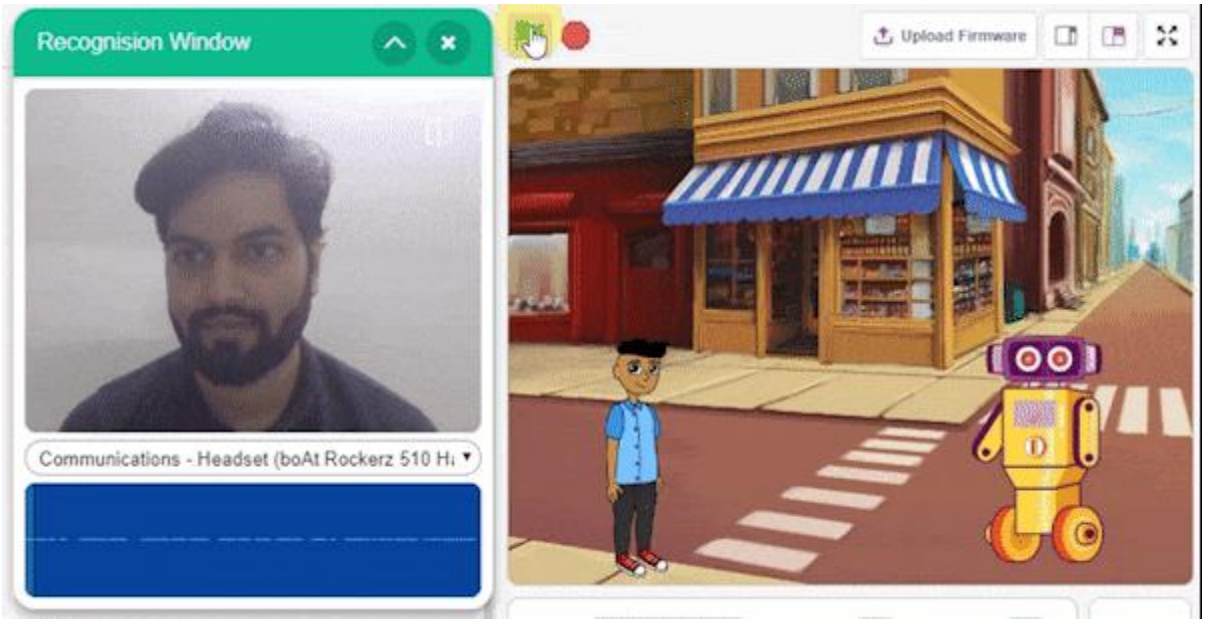
Speech Result Block

To get the result, we have the speech recognition result block. It reports the last text detected from the speech.



Here is a simple example of how to use the speech recognition blocks:





In the next topic, we will make a project using the speech recognition blocks.

(END)