**🚀 How to Use ESP-SR (Espressif Speech Recognition) on ESP32-S3**

ESP-SR is **Espressif’s offline voice recognition framework** that works on **ESP32-S3** for **wake-word detection, keyword spotting, and speech commands**.

**🛠️ Requirements**

🔹 **ESP32-S3 Board** (ESP32-S3 DevKit, ESP32-S3 Box, etc.)  
🔹 **ESP-IDF (Espressif IoT Development Framework)**  
🔹 **Microphone (I2S or PDM, e.g., INMP441, SPH0645)**

**1️⃣ Install ESP-IDF (Required for ESP-SR)**

Since **ESP-SR does not work with Arduino IDE**, you must install **ESP-IDF**.

**🔹 (A) Windows Installation**

1. Download the **ESP-IDF Installer** from [Espressif’s official site](https://github.com/espressif/esp-idf/releases).
2. Install the **ESP32-S3 Toolchain** and **Python dependencies**.
3. Open **ESP-IDF PowerShell** and set up the environment:

. $HOME/esp/esp-idf/export.ps1

**🔹 (B) Linux/macOS Installation**

1. Clone ESP-IDF:

git clone --recursive https://github.com/espressif/esp-idf.git

cd esp-idf

./install.sh

1. Set up the environment:

sh

CopyEdit

. ./export.sh

**2️⃣ Download and Set Up ESP-SR**

1. **Clone ESP-SR Repository**

git clone --recursive https://github.com/espressif/esp-sr.git

cd esp-sr/examples/interactive\_speech\_commands

1. **Set ESP32-S3 as the Target**

idf.py set-target esp32s3

1. **Configure the Speech Model**

idf.py menuconfig

* + Navigate to **Component Config → ESP-SR**
  + Select **Wake Word & Command Model**
  + Choose your **language (English, Chinese, or Custom)**

1. **Compile and Flash to ESP32-S3**

idf.py build flash monitor

This will **upload ESP-SR** to your ESP32-S3 and start listening for commands.

**3️⃣ Connect a Microphone (I2S)**

If using an **I2S microphone** (like INMP441), connect it as follows:

| **Microphone Pin** | **ESP32-S3 Pin** |
| --- | --- |
| **VCC** | **3.3V** |
| **GND** | **GND** |
| **SCK (BCLK)** | **GPIO18** |
| **WS (LRCLK)** | **GPIO19** |
| **SD (DATA)** | **GPIO21** |

**4️⃣ Train Custom Voice Commands (Optional)**

To **add new words (e.g., Korean, custom commands)**:

1. **Record your voice samples** (.wav files).
2. Use **Edge Impulse or TensorFlow** to train a **Keyword Spotting Model**.
3. Convert it to **TFLite format** and replace the existing model in esp-sr/models/.