# Mini guide for Buzz captions with Open AI Whisper

## Introduction

Buzz Captions is a tool for creating audio and video transcripts, powered by OpenAI Whisper. It offers captioning services in multiple languages and provides fairly accurate transcripts for content. In this mini guide, we will provide step-by-step instructions for installing Buzz Captions on Mac and Windows systems, as well as for conducting live recordings and offline file tasks using Buzz Captions. All details about Buzz Captions can be found in this GitHub repository here.

For KU Windows users, to install Buzz Captions, you may need to follow the "Admin by Request" manual, which grants temporary administrator rights when necessary for application installations. This security measure ensures the prevention of accidental installations from sources like phishing emails. For detailed steps, please refer to the step-by-step 'Admin by Request' Manual available here.

## Installation

1. **Mac systems**. Buzz Captions provides two versions. Users can either purchase and install the app directly from the App Store or choose the free open-source version and manually install it. The latest version was updated on March 9th.

- For the mac-native version, buzz captions is currently available for purchase and download at the App Store.
  App store> Search for "Buzz Captions"
- Follow this link to get the latest free version for mac system. Go to Assets > choose and download the "Buzz-0.7.2-mac.dmg" file > open the downloaded file > drag it to your applications folder
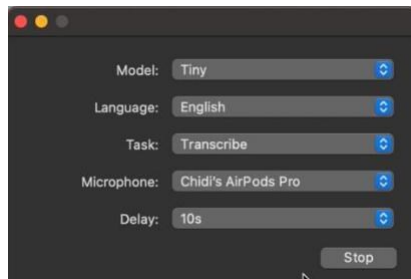
2. **Windows systems**. Follow this link and download the latest version. Go to Assets > download the "Buzz-0.7.2-windows.exe" file > open the downloaded file > select install mode > select destination location > select create a desktop shortcut (if necessary) > hit install.

## Getting Started

Buzz Captions supports both real-time and offline transcription and translation. The user interface for Buzz Captions is simple and user-friendly. While the details of the UI design for the Mac and Windows versions are slightly different, the general ideas are the same. It is divided into two main parts: first, live recording, where you can perform tasks on the played video or directly record via Buzz Captions; and second, importing local files for transcription and translation.

Buzz provides a list of models for users to choose from, ranging from tiny to large. The default setting starts with the tiny model as it runs fast, but you are welcome to choose larger models for your needs. The larger the model you choose, the better the accuracy. For the first-time use of a model, it takes a few minutes to download the model to your device. The Windows version supports Whisper and Hugging Face models, while the Mac version provides more models.
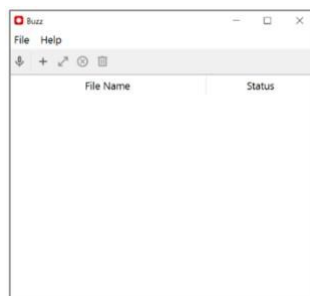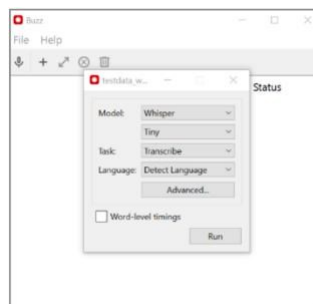
# Getting started with Mac version



You can choose any of the available models, languages, and tasks (translation or transcription). If you already know the language of the content, you can specify it in the language column. Buzz Captions also supports a "detect language" function. Select "Delay and record," and after a time delay, it will translate the speech into text. This process may take longer to run.

To import a local file from your computer, you can find the import option on the top left side of the bar.

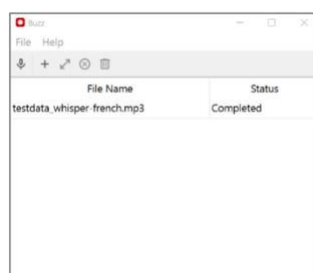# Getting started with Windows version



Windows users open buzz captions, the small window on the left will pop up. You can import a media file by clicking file and select the local file on your computer.



After importing the file, choose the model, task and language. "Word-level timing" is a feature in Buzz Captions that refers to the ability to automatically time the captions to match the audio and video content of a video file. It synchronizes with the audio or video file. It can save time and effort compared to manually timing the captions, and it ensures that the captions are accurately aligned with the spoken words in the video.

Click run.



After running a task, you can check the status. By clicking on the file, you will get the transcript which is ready and available for export in three formats: TXT, VTT, and SRT.

# Some Considerations and Tips

Using Buzz Captions is a simple and straightforward process that can help you with online and offline translation and transcription of audio. However, there are a few things to keep in mind.

- Buzz Captions requires a capable computer to run large media files. We suggest breaking down long audio/video into several shorter segments for efficiency.
- When using in Danish, you might need to choose a larger model, such as 'Medium'. Keep in mind that the larger the Model the higher the quality and longer computation time.
- The performance of live recording is not always stable, and users may encounter slow reactions and buffering. Additionally, the accuracy rate for local media file transcription is generally higher than that for live recording.