

User Manual

CA400 Final Year Project

Project Title:	OCR Text to Speech Reading Aid
Student 1:	Róisín O'Rourke - 19360491
Student 2:	David Weir - 19433086
Supervisor:	Paul Clarke
Date Completed:	05/05/2023

Table of Contents

1. Description	3
2. Installation	3
3. Walkthrough	4
3.1 Upload Files	4
3.2 Edit Text	5
3.3 Modify Text	6
3.4 Convert File Format	7
3.5 Mp3 player	8

1. Description

The application is an Optical Character Recognition (OCR) reading aid. It has the capacity to scan printed and handwritten text, detect the text language, translate the text into a set sample of languages and summarise it and then finally convert the text into an mp3 or PDF format. The text to speech functionality will allow users to convert handwritten documents into pdfs and listen to them in audio format. It is a Python based application that can be installed on the computer or device of the user.

In an educational setting, the system will be able to help students who wish to listen to their notes and study materials as they revise. The translation abilities of the system will also especially aid international students who can scan books/notes not provided in their home language and then listen to them in their own language.

2. Installation

The project can be run by cloning the repository that is available on gitlab. To clone the project, run the following command in the terminal:

Git clone <https://gitlab.com/computing.dcu.ie/weird4/2023-ca400-weird4-orourr28>

All necessary dependencies for the application can be found in the requirements.txt. To install these run the command:

Pip install requirements.txt

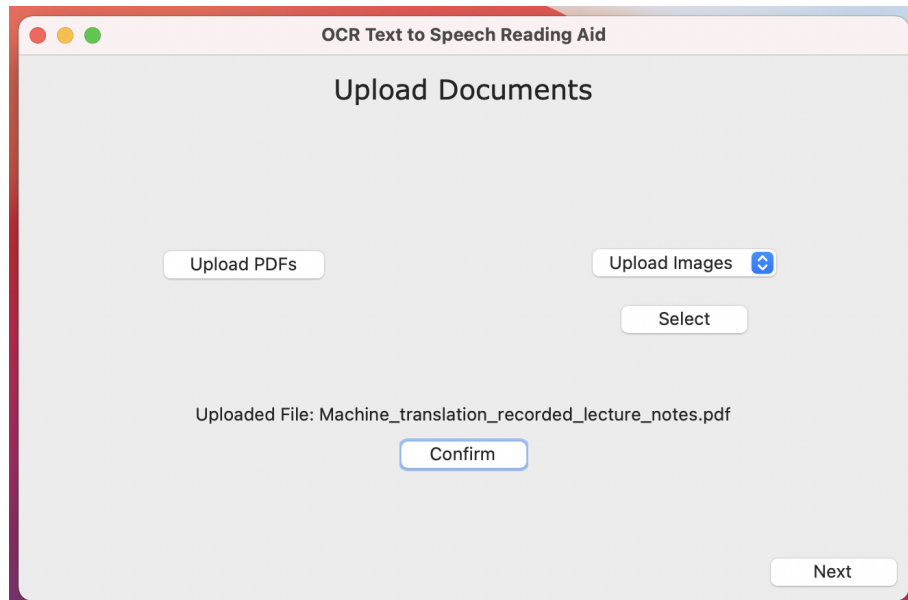
To run the application change directory into the gui folder, which is found under the src folder. Once you have moved into the directory, run the following command:

python gui.py

The application will begin running and the GUI will appear.

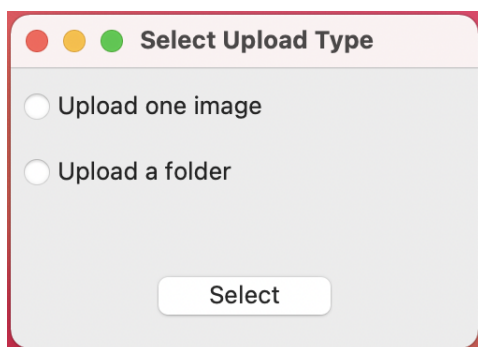
3. Walkthrough

3.1 Upload Files



The first page of the application and the first step is to upload the documents that the user wishes to modify and change the format. The user is given two initial options - to upload a pdf of text, or to upload images of text. If the user wishes to upload a pdf, upon selecting that option, the file explorer will open and the user may select a file.

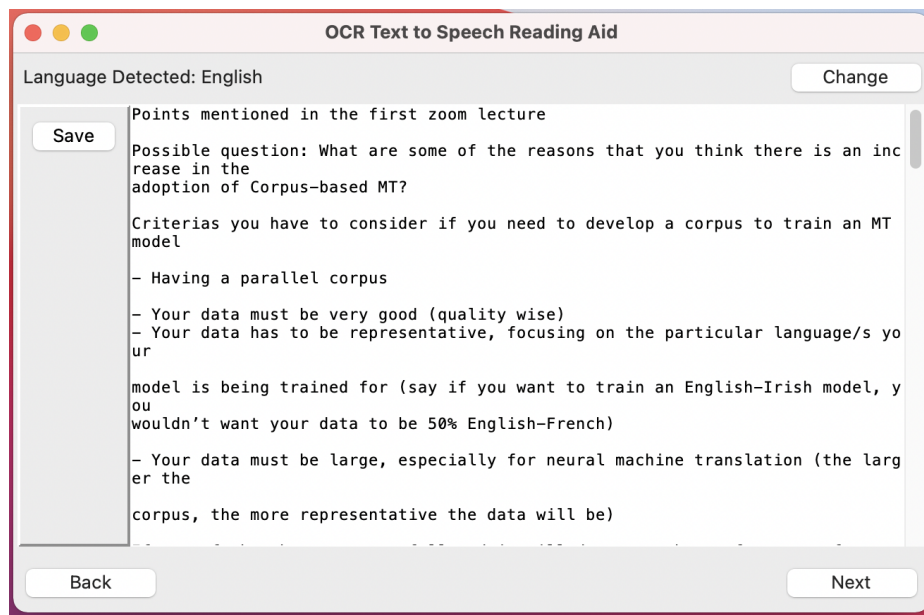
If the user chooses to upload images of text they must click the dropdown menu on the right side of the screen. The two options given are to upload images of printed text or to upload images of handwritten text. If the printed option is selected, a popup window will appear, giving the user the option to choose between uploading a single image or a folder of images.



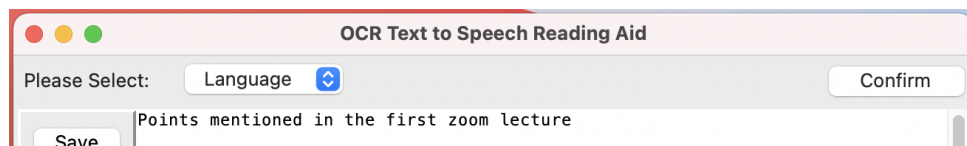
Once an option has been selected and the 'Select' button is clicked, the file explorer will open and the user can upload a file or select a directory, depending on the option chosen.

If the user selects the handwritten text option, the file explorer will open and the user can select either one or multiple images to upload. Once the user is satisfied with their selection, they click confirm, and then the opportunity to move onto the next screen appears.

3.2 Edit Text

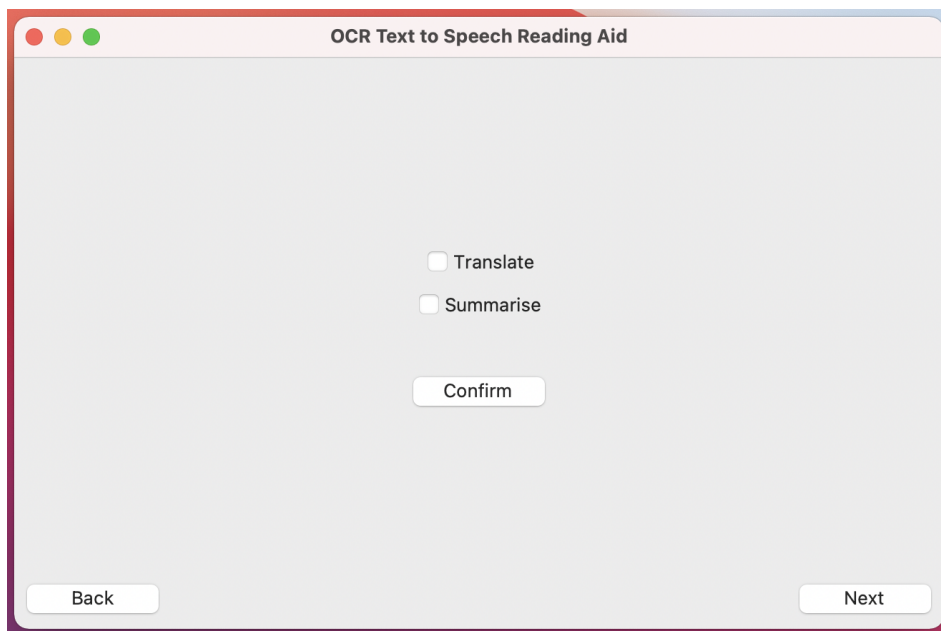


This screen gives users the opportunity to correct any errors or add additional text. The page displays the text extracted from the uploaded documents as well as the detected language of the text. The user can scroll through the text and type in the textbox. Once the user is satisfied with the changes made, they can click the 'Save' button on the left side of the screen to update the text.

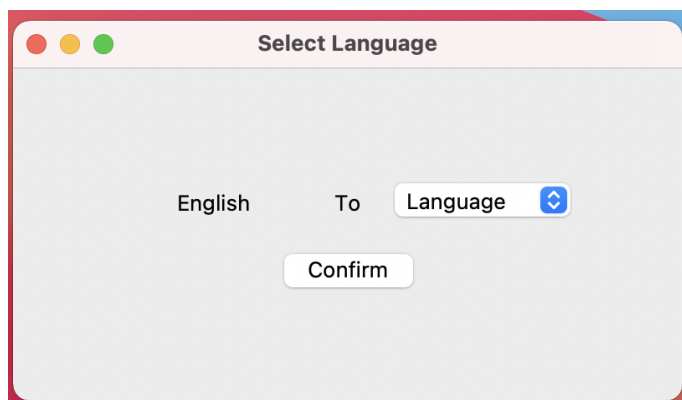


The language detected is displayed at the top of the screen. If the language has been detected incorrectly or if the detector has failed to detect the language, the user can manually select the language. Upon clicking 'Change', a dropdown menu appears with the different languages available and the user can choose the correct language.

3.3 Modify Text

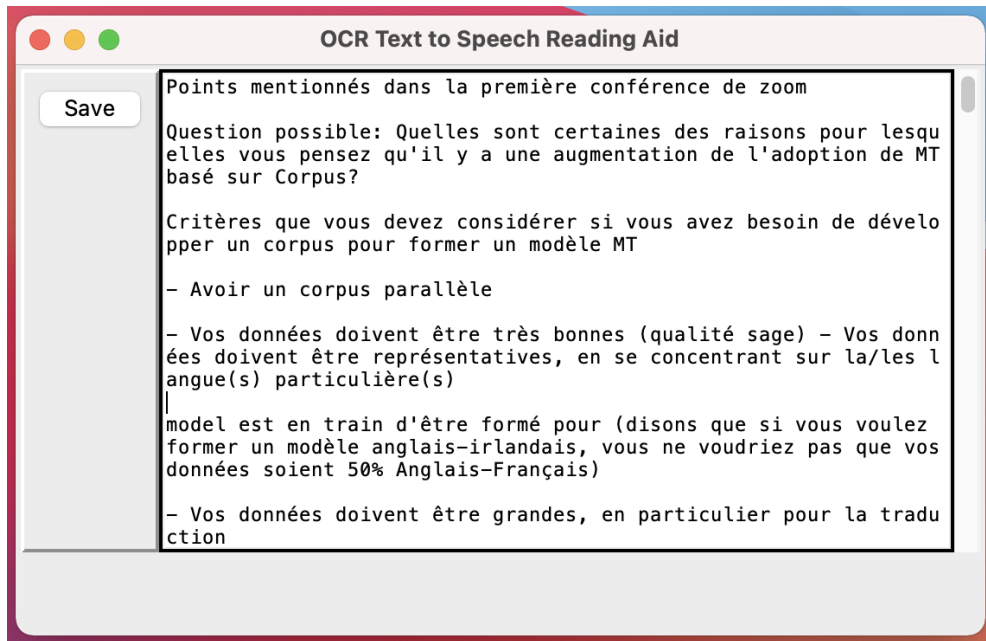


If the user is satisfied with the text and the language, they will progress to the next frame. Here, they have the opportunity to translate and / or summarise the text they have uploaded.



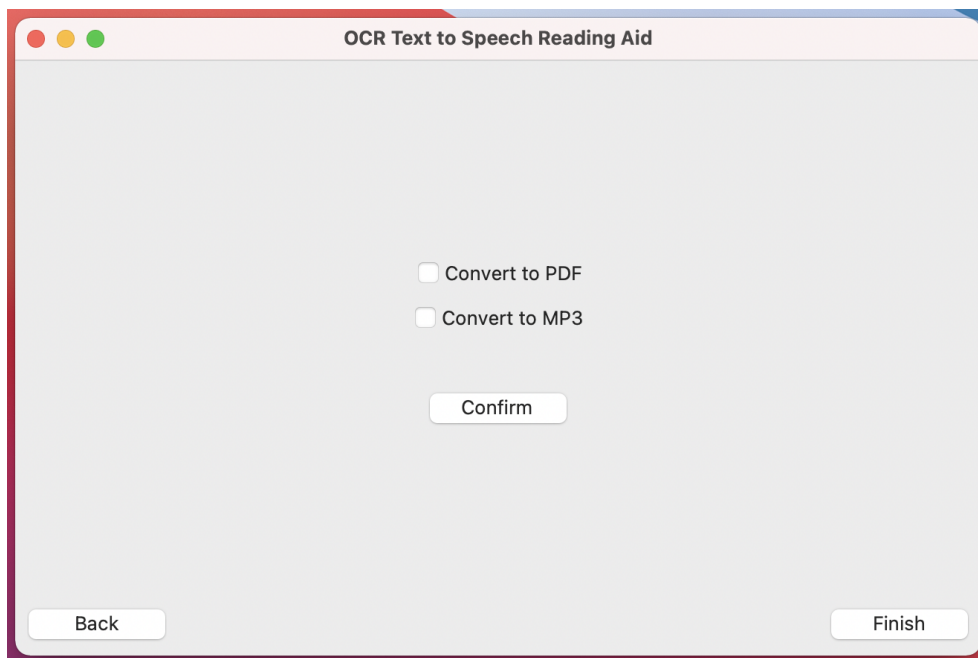
If translation has been selected, a popup window will appear. The user will be asked which language they wish to translate their text into, from the supported languages of English, German and French.

If the user wishes to skip the translation and only summarise, the translation popup will not be displayed and it will simply display the next popup showing the modified text.

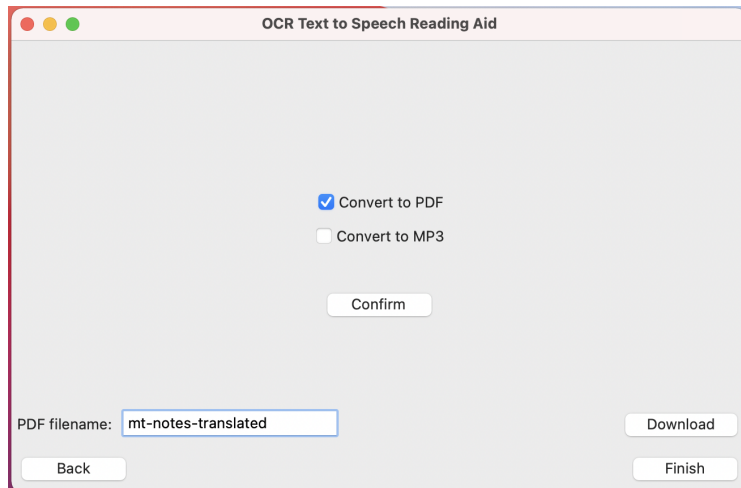


This popup window displays the modified text. The user once again has the opportunity to make any edits to the text and click the 'Save' button to save any changes made. Once the user is satisfied, they can close the popup by clicking on the red 'x' button.

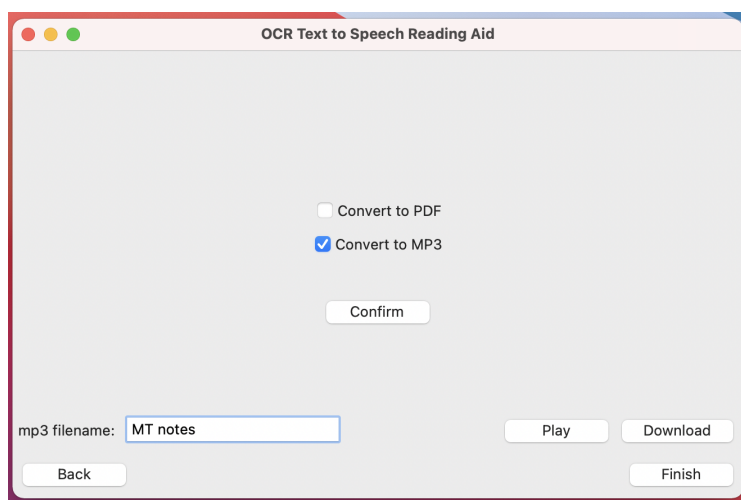
3.4 Convert File Format



The final step of the application allows the users to convert their text into a new format - an mp3 recording or a PDF. The user selects the output options of their preference and then clicks confirm.

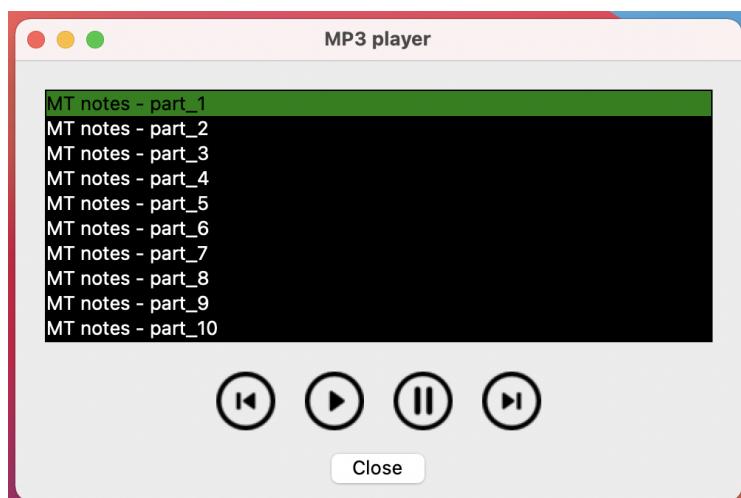


If the user has selected the 'convert to PDF' option, a frame appears with a download button and a text box to enter the name they wish to use for the PDF.



If the user has selected the 'convert to MP3' option. The user now has the option to play the mp3 recording or download. If they choose 'play', the mp3 player will open.

3.5 Mp3 player



The mp3 player is displayed if the user chooses to play the mp3 recording. The mp3 player displays the recording split into smaller sections. The four buttons displayed allow the user to play all the tracks, to pause and unpause, to go back to the previous track and to skip to the next track. When a track

completes playing, the next track will begin automatically.

To close the player, click on the red 'x' button.