



Transcribe RO - GUI Guide

Overview

The Transcribe RO GUI provides a user-friendly graphical interface for audio transcription and translation. It's perfect for users who prefer a visual interface over command-line tools.

Features

File Selection

- **Browse Button:** Opens a file picker to select audio files
- **Supported Formats:** MP3, WAV, M4A, FLAC, OGG, WMA, AAC, OPUS
- **File Path Display:** Shows the full path of the selected file
- **Clear Button:** Removes the current selection

Language Selection

- **15+ Languages:** Choose from Romanian, English, Spanish, French, German, Italian, Portuguese, Russian, Chinese, Japanese, Arabic, Hindi, Dutch, Polish, Turkish
- **Romanian Default:** Romanian is pre-selected as the source language
- **Scrollable List:** All languages are accessible via radio buttons
- **Auto-Detection:** The Whisper model will still auto-detect the language, but your selection helps optimize the process

Settings

Model Size

- **tiny:** Fastest, least accurate (good for testing)
- **base:** Good balance of speed and accuracy (default)
- **small:** Better accuracy, slower
- **medium:** High accuracy, requires more resources
- **large:** Best accuracy, slowest

Device Selection

- **auto:** Automatically detects the best available device (recommended)
- **cpu:** Forces CPU processing
- **mps:** Forces Apple Silicon GPU (M1/M2/M3 Macs)
- **cuda:** Forces NVIDIA GPU

Two-Panel Results Display

Left Panel: Original Transcript

- Displays the transcribed text in the source language
- Scrollable text area with word wrap
- Copy button to copy text to clipboard
- Save button to save transcript to a file

Right Panel: Romanian Translation

- Shows the Romanian translation of the transcript
- Special case: If audio is already in Romanian, displays a message indicating no translation is needed
- Scrollable text area with word wrap
- Copy button to copy translation to clipboard
- Save button to save translation to a file

Progress Tracking

- **Progress Bar:** Animated progress indicator during processing
- **Status Messages:** Real-time updates on current operation
- **Color-Coded Status:**
 - Gray: Ready/Idle
 - Blue: Information
 - Orange: Processing
 - Green: Success
 - Red: Error

Control Buttons

- **Start Transcription:** Begins the transcription process
- **Stop:** Cancels the current operation (with confirmation)

How to Use

Step 1: Launch the GUI

Choose one of the following methods:

Linux/macOS:

```
./run_gui.sh
```

Windows:

```
run_gui.bat
```

macOS (Double-click):

Double-click `run_gui.command` in Finder

Direct Python:

```
python transcribe_ro_gui.py
```

Step 2: Select Audio File

1. Click the “**Browse...**” button
2. Navigate to your audio file
3. Select the file and click “**Open**”

4. The file path will appear in the text field

Step 3: Configure Settings

1. **Model Size:** Select based on your needs (default: base)

- Use “tiny” or “base” for quick results
- Use “medium” or “large” for best accuracy

2. **Device:** Leave as “auto” unless you have specific needs

- “auto” will automatically use GPU if available
- Force “cpu” if you encounter GPU issues

3. **Source Language:** Select the language of your audio

- Romanian is selected by default
- The model will still auto-detect, but your selection helps

Step 4: Start Transcription

1. Click “ Start Transcription”

2. Wait while the model loads (first time may take longer)

3. Monitor progress via the progress bar and status messages

4. Process may take several minutes depending on:

- Audio length
- Model size
- Device (GPU is faster than CPU)

Step 5: Review Results

Once complete:

1. **Left Panel:** Shows original transcript in source language

2. **Right Panel:** Shows Romanian translation (or message if already Romanian)

3. A success message will appear

Step 6: Export Results

For each panel:

-  **Copy:** Copies text to clipboard
-  **Save:** Opens save dialog to save as text file
- Original transcript: <filename>_transcription.txt
- Romanian translation: <filename>_translated_ro.txt

Special Cases

Audio Already in Romanian

If the detected language is Romanian:

- Only the transcript is shown in the left panel
- Right panel displays: “✓ Source audio is already in Romanian. No translation needed.”
- This is by design - no need to translate Romanian to Romanian!

Translation Not Available

If the `deep-translator` library is not installed:

- The GUI will still transcribe the audio
- Translation will not be performed
- Install dependencies: `pip install -r requirements.txt`

Processing Errors

If an error occurs:

1. An error message will be displayed
2. Check the status message for details
3. Common issues:
 - File not found
 - Unsupported format
 - Insufficient memory for large models
 - GPU not available (will fallback to CPU)

Keyboard Shortcuts

- **Ctrl+O** (or **Cmd+O** on Mac): Open file browser (when focused on window)
- **Ctrl+C** (or **Cmd+C**): Copy selected text (when text is selected)
- **Ctrl+A** (or **Cmd+A**): Select all text (in text panels)

Performance Tips

For Faster Processing

1. **Use GPU:** Set device to “auto” to automatically use GPU
 - Apple Silicon: 3-5x faster with MPS
 - NVIDIA GPU: 5-10x faster with CUDA
2. **Choose Smaller Model:** Use “tiny” or “base” for quick results
 - “tiny”: Fastest, good for general use
 - “base”: Good balance (default)
3. **Close Other Applications:** Free up system resources

For Best Accuracy

1. **Use Larger Model:** Select “medium” or “large”
 - Slower but more accurate
 - Better for complex audio or multiple speakers
2. **High-Quality Audio:** Use high-quality audio files
 - Less background noise = better accuracy
 - Clear speech = better results
3. **Correct Language Selection:** Select the correct source language
 - Helps optimize the transcription process

Troubleshooting

GUI Won't Launch

Problem: Script fails to start

Solutions:

1. Make sure Python 3.8+ is installed: `python3 --version`
2. Activate virtual environment: `source venv/bin/activate` (Linux/macOS) or `venv\Scripts\activate` (Windows)

3. Install dependencies: `pip install -r requirements.txt`
4. Check for tkinter: `python3 -m tkinter` (should open a window)

“Module Not Found” Error

Problem: Import errors when starting GUI

Solution:

```
# Make sure you're in the correct directory
cd /path/to/transcribe_ro

# Activate virtual environment
source venv/bin/activate # Linux/macOS
venv\Scripts\activate # Windows

# Install/reinstall dependencies
pip install -r requirements.txt
```

GUI Freezes During Processing

Problem: Window becomes unresponsive

Reason: This is normal! The GUI uses threading to prevent freezing, but initial model loading may cause brief unresponsiveness.

Solutions:

- Wait for the model to load (first time takes longer)
- Don't click repeatedly - be patient
- Check status message for progress updates

Cannot Select File

Problem: File browser doesn't open or crashes

Solutions:

1. Make sure you have read permissions for the audio file
2. Try moving the audio file to a more accessible location (e.g., Desktop)
3. Restart the GUI application

Translation Not Working

Problem: Right panel shows original text instead of translation

Solutions:

1. Check if audio is already in Romanian (no translation needed)
2. Verify `deep-translator` is installed: `pip list | grep deep-translator`
3. Check internet connection (Google Translate API requires internet)
4. Try again - translation service may have temporary issues

GPU Not Being Used

Problem: Status shows “Running on CPU” instead of GPU

Solutions:

1. Verify PyTorch is installed: `python3 -c "import torch; print(torch.__version__)"`
2. For Apple Silicon: Check MPS availability: `python3 -c "import torch; print(torch.backends.mps.is_available())"`
3. For NVIDIA: Check CUDA availability: `python3 -c "import torch; print(torch.cuda.is_available())"`
4. Try forcing device: Set device to “mps” or “cuda” manually

System Requirements

Minimum

- Python 3.8+
- 4GB RAM
- 1GB free disk space
- CPU: Any modern processor

Recommended

- Python 3.9+
- 8GB RAM
- 2GB free disk space
- GPU: Apple Silicon (M1/M2/M3) or NVIDIA GPU with CUDA support

Display

- Minimum resolution: 1024x768
- Recommended: 1280x800 or higher
- The window is resizable to fit your screen

Platform-Specific Notes

Windows

- Use `run_gui.bat` to launch
- FFmpeg must be installed and in PATH
- Windows Defender may scan the first run (this is normal)

macOS

- Use `run_gui.command` (double-click) or `./run_gui.sh`
- Grant microphone permissions if prompted (not required unless recording)
- Apple Silicon Macs will automatically use MPS GPU acceleration

Linux

- Use `./run_gui.sh` to launch
- Install tkinter if needed: `sudo apt-get install python3-tk` (Ubuntu/Debian)
- Grant execute permissions if needed: `chmod +x run_gui.sh`

Advanced Features

Resizable Window

- Drag window edges to resize
- Text panels automatically adjust to window size
- Minimum size enforced to maintain usability

Clipboard Integration

- Copy button copies entire text to clipboard
- Use Ctrl+C/Cmd+C for selected text
- Paste anywhere with Ctrl+V/Cmd+V

File Saving

- Save dialogs remember last location
- Default filenames based on original audio file
- UTF-8 encoding ensures proper display of Romanian diacritics

Support

For issues or questions:

1. Check this guide first
2. Review the main [README.md](#) (README.md)
3. Check [TROUBLESHOOTING.md](#) (TROUBLESHOOTING.md) if available
4. Open an issue on GitHub

Version Information

- **GUI Version:** 1.0.0
 - **CLI Version:** 1.1.0
 - **Last Updated:** January 2026
-

Enjoy using Transcribe RO GUI! 🎉