

How to Verify Dual-File Output

Quick Verification Steps

Step 1: Run Unit Tests

```
cd /home/ubuntu/transcribe_ro
python test_dual_file_output.py
```

Expected Output:

```
✓ ALL TESTS PASSED!
```

Step 2: Test with Real Audio File

If you have an audio file, test the dual-file output:

```
# Test with translation (creates 2 files)
python transcribe_ro.py your_audio.m4a --debug
```

What to Look For:

1. During Processing (in debug output):

```

'''
=====
STEP: PREPARE OUTPUT
=====
Original output path: ../your_audio_transcription.txt
Translated output path: ../your_audio_translated_ro.txt
=====
STEP: WRITE ORIGINAL TRANSCRIPTION FILE
=====
Writing to: ../your_audio_transcription.txt
✓ Original transcription saved to: ../your_audio_transcription.txt
=====
STEP: WRITE TRANSLATED FILE
=====
Writing to: ../your_audio_translated_ro.txt
✓ Romanian translation saved to: ../your_audio_translated_ro.txt
'''

```

1. In Final Summary:

```
=====
PROCESSING COMPLETED SUCCESSFULLY!
=====
```

```
Detected language: en
```

```
Original transcription: your_audio_transcription.txt
Romanian translation: your_audio_translated_ro.txt
✓ Two files created: original transcription + Romanian translation
```

2. File System Check:

```
bash
ls -la your_audio*
```

Should show:

```
your_audio.m4a
your_audio_transcription.txt  <- Original transcription
your_audio_translated_ro.txt  <- Romanian translation
```

Step 3: Verify File Contents

Check Original File:

```
cat your_audio_transcription.txt
```

Should see:

- Header: "TRANSCRIPTION RESULTS (ORIGINAL LANGUAGE)"
- Original language text only
- NO Romanian translation section

Check Translated File:

```
cat your_audio_translated_ro.txt
```

Should see:

- Header: "ROMANIAN TRANSLATION"
- Romanian translated text
- Metadata showing original language

Step 4: Test All Formats

```
# Test JSON format
python transcribe_ro.py your_audio.m4a --format json --debug

# Test SRT format
python transcribe_ro.py your_audio.m4a --format srt --debug

# Test VTT format
python transcribe_ro.py your_audio.m4a --format vtt --debug
```

Each should create TWO files with appropriate suffixes.

Step 5: Test No-Translation Mode

```
# Should create only ONE file
python transcribe_ro.py your_audio.m4a --no-translate
```

Should create only:

- your_audio_transcription.txt (no translated file)

Expected File Naming

Input File	Format	Original File	Translated File
audio.m4a	txt	au-dio_transcription.txt	au-dio_translated_ro.txt
audio.m4a	json	au-dio_transcription.json	au-dio_translated_ro.json
audio.m4a	srt	au-dio_transcription.srt	au-dio_translated_ro.srt
audio.m4a	vtt	au-dio_transcription.vtt	au-dio_translated_ro.vtt

Troubleshooting

Issue: Only one file is created

Possible Causes:

1. Audio is already in Romanian (no translation needed)
 - Check debug output: "Audio is already in Romanian. No translation needed."
1. `--no-translate` flag was used
 - Remove the flag to enable translation
2. Translation failed
 - Check debug output for translation errors
 - Verify internet connection (translation requires online access)

Issue: Files not found

Solution:

```
# List all files in current directory
ls -la

# Search for transcription files
find . -name "*transcription*" -o -name "*translated_ro*"
```

Issue: Debug output not showing both files

Solution:

- Make sure you're using the updated `transcribe_ro.py`
- Run with `--debug` flag
- Check that translation is actually happening (not Romanian audio)

Success Indicators

✔ **Verification Successful** if you see:

1. TWO separate files created
2. Original file contains only original language text
3. Translated file contains only Romanian translation
4. Both files have correct naming convention
5. Debug output shows both file paths
6. Final summary mentions both files
7. All tests pass

Quick Test Command

Run this complete test sequence:

```
# Run unit tests
python test_dual_file_output.py

# If you have audio file, test real transcription
python transcribe_ro.py your_audio.m4a --debug

# Verify files were created
ls -la *transcription* *translated_ro*

# Check first few lines of each file
head -20 *transcription.txt
echo "---"
head -20 *translated_ro.txt
```

Need Help?

If verification fails:

1. Check the debug output carefully
2. Verify all changes were applied to `transcribe_ro.py`
3. Ensure you're running the correct script version
4. Check file permissions
5. Look for error messages in the output

Last Updated: 2026-01-12

Verified Version: transcribe_ro.py v1.1.0