

Breakout 2: TRANSCRIPTION

FILES

english_audio.mp3
turkish_audio.mp3
portuguese_audio.mp3
spanish_audio.mp3

We will try transcribing audio recorded in different languages using different online services; note that not all languages will be available in all services.

STEP 1. Sign up for an account

Sign up for accounts with the following transcription services. You can cancel the accounts if you think you won't need them afterwards. Additional details and links to tutorials are available on the *Resources* page.

1. [Otter.ai](#).
2. [Happy Scribe](#)
3. [Watson](#); note the “detect multiple speakers” option.

STEP 2: Transcribe

Transcribe the following audio files with the different services (if possible). Save the resulting transcriptions files as Google Docs or .txt files. Use the following naming convention: transcription_[language]_[name of service]

1. spanish_audio.mp3
2. indonesian_audio.mp3
3. mandarin2_audio.mp3

STEP 3 (IF TIME PERMITS): Translate

1. Try translating the files with Google Docs
2. File names: translation_[language]_[name of service]
3. Paste the translations and the original text (from the .txt or .docx files) into the Text Comparison Website <https://www.diffchecker.com/diff>. You may have to adjust the spacing in the text files to make the text align.

STEP 4: Discuss

1. Do the transcriptions appear accurate?

2. What about the resulting translations? Is there a loss in quality when going from audio-to-text-to-English when compared with text-to-English
3. What would be the best approach for research:
 - a. transcribing, editing the transcription, then translating, or
 - b. transcribing, translating, then correcting?