Breakout 2: TRANSCRIPTION

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

1. Download the following audio files. Note that the link takes you to the page where the file is archived; click the 'download' button to save it to your computer.

    - `turkish_audio.mp3`

    - `spanish_audio1_voice.mp3`

2. Now transcribe these audio files with each of the different services (note that not all services will transcribe all languages).

3. Save the resulting transcriptions files as Google Docs or .txt files. Use the following naming convention: `transcription_[language]_[name of service]`

STEP 3: Translate

1. Translating these transcriptions to English with Google Docs

2. Save with the following file names: `translation_[language]_[name of transcription service]` (e.g., `translation_spanish_happyscribe`)

3. Paste the *translations from transcription* and the *translation from the original .txt files* (used in the translation breakout) into the Text Comparison Website. 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 translations from transcriptions? Is there a loss in quality when going from audio-to-text-to-English when compared with text-to-English

Last update 27 March 2023

3. What would be the best approach for research:

    a. transcribing, editing the transcription, then translating, or
    b. transcribing, translating, then correcting?