Notes on epub to audiobook

June 6, 2024

John Nash
for
Virtual Meeting of Linux-Ottawa

Motivations

- I have about a dozen novels / collections as epub files on obooko.com site (free download for private individuals)

- Audiobooks becoming popular. Format m4b seems to be the current popular form, but older players may need mp3.

- Can convert m4b to mp3 with **ffmpeg**

- Recording the book by voice is tedious and requires careful attention to ambient conditions, good microphone and other details

- Can we use **tts** tools (Text to speech)?

Issues

- TTS software requires a large infrastructure

- Details, details, details

- Systems often set up for one sentence at a time

- Multiple languages (English main, French, Dutch, ...).

- Proper names of people and places

- Dialect: "'arry went to 'ampstead. 'arry lost 'is 'at. 'arry's mother said to 'arry, 'arry where's your 'at?"

My tries

- Primarily based on eput2tts and epub3tts-edge by Christopher Aedo

- https://github.com/aedocw/

- First uses Coqui AI TTS, OpenAI or MS Edge

- Latter just MS Edge TTS

epub2tts-edge

- Simple command

  epub2tts-edge mybook.epub

  this quickly creates mybook.txt, then run

  epub2tts-edge mybook.txt

  which takes a LONG time (I don't have CUDA capability), about 1/2 hour

  Generally want to add speaker and cover options

  Result surprisingly good, but ...

- Pronunciation may depend on context, and quite often incorrect ("read" present/past)

- Has possible multilingual capability with some "speakers", but changes voice e.g., in French

- I haven't worked out all the controls

- Speakers in form en-US-AndrewNeural (multilingual), en-GB-ThomasNeural (English only)

epub2tts-edge

- Can adjust mybook.txt with dividers in form

  # Part 24

- Can change text to "sound" better. I used this quite a lot for a book **Leftovers**

  - Paule --> Pohl

  - Sault Ste Marie --> Soo Saint Marie

    tricky!!!

- Punctuation: Mr., Mrs. Ms. --> unwanted pauses

  - Mister and Missus and Miz help. (Ms may be spelled out).

epub2tts

- Simple command, but speakers now coded e.g. p335 (female, default), p340 is somewhat British voice I decided to use. None fully what I would like.

- I could not get '--xtts' option working to use voice samples (i.e., my own voice)

- Builds on the underlying "tts" command, but I found limited documentation of the VERY busy syntax.

epub2tts

- Seems to work in single pass on either epub or txt file. If txt with '# Part ???' dividers, creates separate .wav files.

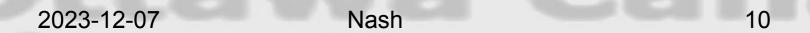- Can use ffmpeg to convert wav to mp3, and mp3wrap to consolidate to fewer files.

Choices?

- Still trying to decide which tool I prefer.

- Both are "almost" there in terms of giving usable audiobooks.

- Useful to convert m4b to mp3 for older players.

- Also useful to "slow" (which also lowers voice pitch) of play. I used Audacity on the mp3. Possibly other ways to do this with m4b

- URLs and email addresses best converted,

  nashjc@ncf.ca to "n a s h j c   at n c f dot c a"

DEMOS

Demo files

Demos?

Leftovers.epub

Leftovers-en-GB-ThomasNeural.m4b

Leftovers-partA.mp3

240603leftovernotes.txt