TTS:

Dataset requirement (for training):

1. .wav files are required for the data set.
2. transcription of the.wav file
3. .interval file for indicating all starts and stops of pronunciation
4. a .txt file for mapping of all resources

```
000001        卡尔普#2陪外孙#1玩滑梯#4。
      ka2 er2 pu3 pei2 wai4 sun1 wan2 hua2 ti1
000002        假语村言#2别再#1拥抱我#4。
      jia2 yu3 cun1 yan2 bie2 zai4 yong1 bao4 wo3
000003        宝马#1配挂#1跛骡鞍#3，貂蝉#1怨枕#2董翁榻#4。
      bao2 ma3 pei4 gua4 bo3 luo2 an1 diao1 chan2 yuan4 zhen3 dong3 weng1 ta4
000004        邓小平#2与#1撒切尔#2会晤#4。
      deng4 xiao3 ping2 yu3 sa4 qie4 er3 hui4 wu4
000005        老虎#1幼崽#2与#1宠物犬#1玩耍#4。
      lao2 hu3 you4 zai3 yu2 chong3 wu4 quan3 wan2 shua3
000006        身长#2约#1五尺#1二寸#1五分#2或#1以上#4。
      shen1 chang2 yue1 wu2 chi3 er4 cun4 wu3 fen1 huo4 yi3 shang4
000007        赵荻#2约#1曹云腾#2去#1鬼屋#4。
      zhao4 di2 yue1 cao2 yun2 teng2 qu4 gui3 wu1
```

Figure 1. capture of '000001-010000.txt',

```
000001.interval    new 1
 1   File type = "ooTextFile"
 2   Object class = "TextGrid"
 3
 4   0
 5   2.66
 6   <exists>
 7   1
 8   "IntervalTier"
 9   "000001.interval"
10   0
11   2.66        audio length
12   17          numb of interval
13   0           start of 'si1'
14   0.27958612055419324     end of 'si1' (in seconds)
15   "sil"       interval word
16   0.27958612055419324
17   0.406291189066745
18   "k"
19   0.406291189066745
20   0.5165142974475827
21   "a2"
22   0.5165142974475827
23   0.6143759357296349
24   "er2"
25   0.6143759357296349
26   0.743141249258651
27   "p"
28   0.743141249258651
29   0.8997198705099344
30   "u3"
31   0.8997198705099344
32   0.9749188136108798
33   "p"
```

Figure 2. an example of an interval file

Data-preprocess:

      Add padding to ensure all input share equal length

      Recording true length (all non-padding value)

      Convert all .wav file to numpy expression

Train:

      Provide the data, [x:sentence label,y:numpy list].

Evaluate:

      See example in the Colab notebook

Colab Notebook

https://colab.research.google.com/drive/1dEZtktOnY-QgIJc7XinzrXrPkyIKqx-E?usp=sharing

(if you get access denied from reaching the notebook, just open it in incognito mode)

For the environment used,

You may check it with

!pip list

In the notebook

For Cantonese version:

Same requirement for the dataset:

      .wav

      .interval with timestamp

      .txt for mapping

Potential source for dataset:

http://compling.hss.ntu.edu.sg/hkcancor/

https://github.com/liesenf/MYCanCor