Installing and Using the Penn Forced Aligner (P2FA) Chinese

Cong Zhang

If you are reading this article, you probably already know what P2FA is – in brief, it is a tool that you can use to align your transcriptions to your audio files, phoneme by phoneme.

It is not the only existing forced aligner. Kelley and Tucker (2018) has a brief summary on this topic:

> " A number of forced aligners already exist. Some rely on hidden Markov models, such as the Penn Forced Aligner [2], the ProsodyLab Aligner [3], MAUS [4], and the Montreal Forced Aligner [1], while Gentle [5] relies on neural networks."

> [1] M. McAuliffe, M. Socolof, S. Mihuc, M. Wagner, and M. Sonderegger, "Montreal Forced Aligner: Trainable text-speech alignment using Kaldi," in Interspeech 2017, 2017, pp. 498–502.

> [2] J. Yuan and M. Liberman, "Speaker identification on the SCOTUS corpus," Journal of the Acoustical Society of America, vol. 123, no. 5, p. 3878, 2008.

> [3] K. Gorman, J. Howell, and M. Wagner, "Prosodylab-aligner: A tool for forced alignment of laboratory speech," Canadian Acoustics, vol. 39, no. 3, pp. 192–193, 2011.

> [4] T. Kisler, F. Schiel, and H. Sloetjes, "Signal processing via web services: The use case WebMAUS," in Digital Humanities Conference 2012, 2012, pp. 30–34.

> [5] R. M. Ochshorn and M. Hawkins, "Gentle," 2017. [Online]. Available: https://lowerquality.com/gentle/

I haver also tried Easy Align (http://latlcui.unige.ch/phonetique/easyalign/easyalign_unpublished.pdf) and a few others, including Praat's own aligner.

To be honest, none of them is accurate. So, please do not skip manual correction even if you have a lot of data.

Five basic steps are involved in installing and using the Mandarin Chinese version of P2FA:
- Installing Cygwin
- Installing htk
- Installing sox
- Installing python
- Downloading the English and the Chinese versions of P2FA

You also need:
Wav sound files (Step 27)
Text files (Step 28)

Steps:

1. Install <mark>Python</mark> https://www.python.org/downloads/ . Version 2.7 is recommended, but I have Python 3 and it works too.

2. Install Cygwin. Under the "Installing Cygwin" section on their homepage, there are choices for a 64-bit installation and a 32-bit installation. Choose the 32-bit one regardless. It has to be **32-bit Cygwin** even if your computer is 64-bit. Here's a link (https://cygwin.com/setup-x86.exe) for the download.

   Install with the following options:
   ```
   python 2
   gcc-g++
   libx11-devel
   automake
   automake1.15
   make
   nano
   ```

   Install it directly in your C drive. I know it makes your drives messy but let's just do this to avoid complications. So, now you have a `C:\cygwin` folder.

3. Run Cygwin

4. Download **htk-3.4.tar.gz** from http://htk.eng.cam.ac.uk/docs/instnix.shtml
   Note that it has to be htk-3.4 (not 3.41 or anything later).

5. Create folder in C: (using windows filesystem. Just do it the usual way - nothing fancy about this step) called **cygwinfolders**. Now you have a folder `C:\cygwinfolders`

6. Copy the htk file to the newly created folder `C:\cygwinfolders`.

7. In the Cygwin console (the one with black background), type:
   `tar xzf htk-3.4.tar.gz`
   Press the Enter key

8. That will create a folder: htk. Now you have `C:\cygwinfolders\htk`

9. Using windows file explorer:
   copy `c:\cygwin\usr\share\automake.1.15\config.guess` to the `C:\cygwinfolders\htk` folder (overwriting the existing file)

10. Change the directory to the htk folder in Cygwin. Type in the Cygwin console:
    `cd htk`
    Press the Enter key
    `./configure`
    Press the Enter key

11. Expand the width of the Cygwin window by dragging the right-hand edge to the right. In Cygwin console, type:
    `nano HLMTools/makefile`

12. A black text editor box will appear: Go down to mkinstalldir: the line below that has spaces in front of "if" delete the spaces and press tab once.

13. Then press:
    `ctrl + x`
    `y`
    Press the Enter key
    And the editor box will close

14. In Cygwin, type:
    `make all`
    Press the Enter key
    `make install`
    Press the Enter key

15. Create a folder called exes in cygwinfolders. Now you have
    `C:\cygwinfolders\exes`

16. Copy all the .exe files from `hlmtools` and `htktools` to the
    `C:\cygwinfolders\exes` folder

17. Download and install sox http://sox.sourceforge.net/

18. Find the sox.exe in
    `c:\program files (x86)\sox (whatever version)`
    and copy it to `c:\cygwin\bin\`

19. In Cygwin, type:
    `nano align.py`

20. Change line 46 to **'rate'** from **'polyphase'**

21. Edit path in <u>environment variables</u> in <u>system properties - Advanced</u> and add link to sox, to htk folder and cygwindfolders folder. This is how:
    - Type environment variables in the windows start box (or Cortana) and press enter when it has found something. (Google "environment variables Windows 10, if you can't find it.)
    - It should then open up the system properties box.
    - Click environment variables at the bottom
    - In the box that opens look for the **'path'** variable and double click it.
    - Then add the folders in the same format as you can see:
      - your sox folder (probably something like `C:\program files(x86)\sox-14-4-2`)
      - `C:\cygwinfolders`
      - `C:\cygwinfolders\exes`

22. Download the aligner for **English**: https://web.sas.upenn.edu/phonetics-lab/files/2016/05/Penn-Phonetics-forced-aligner-2d2jfwb.zip

23. Download the aligner for **Chinese**: http://languagelog.ldc.upenn.edu/myl/ChineseAligner.tgz

24. Unzip each of them into their own folders. Now we have `C:\cygwinfolders\p2fa` and `C:\cygwinfolders\cp2fa` (Actually, I use `ppfa` instead of `p2fa`, just in case numbers cause problems.)

25. Copy align.py from the English version `C:\cygwinfolders\p2fa` to the Chinese version `C:\cygwinfolders\cp2fa` This is because on my computer, the Chinese version does not create proper TextGrid files. The code seems to be incomplete for some reason. You can try `Calign.py` in the Chinese version to see whether it works. Make sure the homedir is changed to cp2fa folder:
    `homedir = 'c:/cygwinfolders/cp2fa'`

26. To see whether it works on the test files, in Cygwin, type:
    `cd C:\cygwinfolders`
    `cd cp2fa`

27. Prepare your files: A .wav file and a .txt file. Make sure the wav file is sampled at 8000Hz or 16000Hz. If your sampling rate is neither of these two, your first option is to modify the `align.py` script:
    Edit line 46:

    `os.system("c:/cygwin/bin/sox.exe " + orig_wav + " -r 8000 " + out_wav + " rate " + soxopts)`
    The spaces inside the quotation marks are very important. Do not ignore.

    The '8000' here can also be 16000.
    Then, replace the 11025 on line 40 with 8000 or 16000 (keep it consistent with the previous rate), depending on your preference. Note the Chinese version does not have a 11025Hz option. Do not change the 11025 occurances.

    Alternatively, you can also resample your files using sox in the command line:
    - At windows start menu, open up cmd (aka Command Prompt)
    - You should then be able to enter:
      `Sox yourfilename.wav -r 8000 yourresampledfilename.wav`

    Or, if these both sound too complicated, just do the resampling in Praat, Audacity, or other software applications that you're comfortable with.

28. The txt file should be your transcription. Make sure the encoding of the txt files is **UTF-8**. It is a good idea to include an 'sp' at each end of your utterance and have spaces between characters.

29. In Cygwin, type:
```
python align.py test/test_8000.wav test/ test_8000.txt test/
test_8000.Textgrid
```
Press the Enter key

The format is "python align.py test/source file.wav textfile.text output file.Textgrid"

This is an example file in the aligner. You can find the TextGrid file in the `C:\cygwinfolders\cp2fa\test` folder now.

There are other tutorials online that are very useful. For example,
Will Styler, Using the HTK and the Penn Phonetics Lab Forced Aligner on Mac OS X, http://wstyler.ucsd.edu/posts/p2fa_mac.html

The one I originally followed was "Using P2FA" by John M. Riebold. It was quite clear but it is not available online anymore.

Acknowledgements:
This guide is based on the guidance notes that Mr Robert Greenock, our IT officer at University of Oxford, wrote for me. Without his guidance of which versions of the software applications to install, I would be stuck at the very beginning, just like last year, and the year before last year, and the year before the year before last year, and the year before the year before the year before last year...