How to make your own custom Kobo dictionaries

Pre-Requisites

Software you will need

- Python
- Penelope: http://www.albertopettarin.it/penelope.html
- Marisa-Trie: http://code.google.com/p/marisa-trie
  - for windows you should find pre-compiled binaries somewhere on the internet (e.g. here: http://www.mobileread.com/forums/showthread.php?t=196925)
  - for linux/mac compile it as noted here: http://code.google.com/p/marisa-trie/wiki/HowTo?wl=en
    - Download the source
    - Run `$ ./configure`
    - Run `$ make`
    - Run `$ sudo make install`
  - Adjust the penelope.py file around line 50 to the appropriate binaries. (On my ubuntu both are in `/usr/local/bin/`.)
- (optional, if using Babylon) PyGlossary: http://sourceforge.net/projects/pyglossary

Steps

1. Decide which dictionary to use. Penelope supports StarDict, XML, CSV and Bookeen Odyssey as input formats. Typically you will have StarDict or CSV. Recommendations are: Babylon and/or dict.cc

2. Download the dictionary

   1. For Babylon (Windows required):
      1. Go to http://www.babylon.com/free-dictionaries/languages/, choose your language and translation, and download the `Babylon10_setup_dict.exe`.
      2. Install it as usual.
      3. It should download your selected dictionary after installation.
      4. Locate the BGL file: Open an explorer and go to `C:\Users\<Username>\AppData\Local\Babylon` (or whereever your user folder is located). There should be a file called `Babylon_FROM_TO.BGL`. You can either copy it somewhere else

or let it stay there, doesn't really matter - if you are an expert.

5. Make sure **Python** & **PyGlossary** is installed / downloaded
6. Convert it to StarDict format: http://code.google.com/p/stardict-3/wiki/ConvertBabylon

   1. make sure **glade2** is installed, on Ubuntu just type in `sudo apt-get install python-glade2` (yes I am using both Windows and Ubuntu for this)
   2. in a console type: `$ ./pyglossary.sh` or `C..> pyglossary.bat`
   3. Navigate to Preferences tab. In the Convert group uncheck "Remove tags after loading". Close the window.
   4. Now execute the following (substitute environment if you are on windows) `$ ./pyglossary.sh (PATH_TO)Babylon_FILENAME.bgl (OUTPUT_PATH)Babylon_FILENAME.ifo`
   5. Note that the StarDict files will be written to the `OUTPUT_PATH` (which is of course the working directory if left empty)
   6. You should have a .ifo, .dict and .idx file now

2. For dict.cc:

   1. Go to http://www1.dict.cc/translation_file_request.php
   2. Let them send the file to your email
   3. You will get a .zip file with a .txt file in it
   4. This file qualifies as a CSV file (with "comma" being whitespace)
   5. Optional: Convert CSV to XML format

3. Convert to Kobo format.

   1. Run the following command for Babylon:

      ```
      $ python penelope.py --sd -p (PATH_TO)Babylon_FILENAME -f FROM_LANGUAGE -to TO_LANGUAGE --output-kobo
      ```

      Note to not add any file extension after FILENAME. Obviously, you will want to replace the ALL_CAPS stuff, with language codes being en, de, fr, it, etc. Optionally, include de `-d` parameter if you run into trouble

   2. Run the following command for dict.cc:

      ```
      $ python penelope.py --csv -p (PATH_TO)dictcc_FILENAME -f FROM_LANGUAGE -to TO_LANGUAGE --output-kobo -d
      ```

Note to not add any file extension after FILENAME. Obviously, you will want to replace the ALL_CAPS stuff, with language codes being en, de, fr, it, etc. Optionally, include de `-d` parameter if you run into trouble (which is the debug mode)

4. You will have `dicthtml_FROM_TO.zip` files which need to be copied on your device to `.kobo/dict/`. Make sure that you name them after an already existing dictionary.