

Several years ago there was a discussion on the FLEx list if there was a way to make consistent changes throughout an entire database, including lexical entries, word analyses and words in texts. Apparently the Bulk Edit in the lexicon will not change word analyses and words in texts.

The conclusion was you needed to do this by manually editing the FWData file, which could be done with the Consistent Changes program. I've also seen it can be done with regular expressions in the Notepad++ program.

Sample data: the Sena 3 project from the FLEx sample databases page:

<https://software.sil.org/fieldworks/download/sample-projects/>

My test orthography change: Throughout the Sena text, change “kh” to “k” (Latin\_Small\_Letter\_K\_with\_Hook).

Preparation:

1. install the data. Download the backup file from the Sample data page. Restore from a backup into a FLEx project.
2. Preserve a copy of the data so I can recover if my edits produce an unreadable version. Copy the **Sena 3.fwdata** file from the FLEx directory (usually *c:\ProgramData\SIL\Fieldworks\Projects\Sena 2*) to somewhere else on my hard drive.

What to change

The fwdata file is in XML format, where there are XML markers inside of < and > characters, and there is text in the vernacular writing system and the analysis writing system. I only want to make changes in the vernacular writing system. This is a part of that file, showing one word analysis that we will want to change

```
<rt class="CmDomainQ" guid="861eba52-7d57-4991-a9be-4c4f2b94a4e5" ownerguid="9d6cbe74-93d6-41fb-b7e2-cc30adb28187">
<ExampleWords>
<AUni ws="en">come up, rise</AUni>
</ExampleWords>
<Question>
<AUni ws="en">(7) What words refer to the sun or moon moving up in the sky?</AUni>
</Question>
</rt>
<rt class="StTxtPara" guid="861f68c4-eb7e-44a3-a1cf-797f615568a9" ownerguid="9be01c79-6ef0-4eee-9999-c1144e91fcc7">
<Contents>
<Str>
<Run ws="seh">akhabweka</Run>
</Str>
</Contents>
<ParseIsCurrent val="True" />
<Segments>
<objsur guid="25e67699-6143-4557-9b8f-fe56745a32f9" t="o" />
</Segments>
</rt>
```

We want to apply changes only to text in the Sena writing system. This is marked by the code **ws="seh"** inside the preceding marker. So we can tell our change process to look for “kh” in any string of text between **ws="seh">** and **<** (the **<** is the beginning of a new XML marker).

How to make the changes:

For both options, close FLEx so you can access the database file:

Option 1: The consistent changes (CC) program. We can write a consistent changes file like this that applies changes to the range of text between the code signifying Sena text, and the beginning of the next XML marker.

```
c convert orthography of Sena sample FLex database
c kh --> k

begin > use(main)

group(main)

'ws="seh">' > dup use(SenaText)

group(SenaText)
c add our orthography changes here

'kh' > 'k'
'Kh' > 'K'

'<' > dup use(main)
```

The group() statements restrict the orthography changes to only Sena text. To be thorough, I added in a rule to enter the upper case K with hook, where Kh is found, but I don't believe this occurs in the data.

You cannot use CC to modify a file directly, it always makes a new output file, processing the changes in that file. So you would have to run CC to create a modified output file, then remove the original **Sena 3.fwdata** file and rename your output file to Sena 3.fwdata

Option 2: a regular expression in Notepad++

You can open the **Sena 3.fwdata** file in Notepad++, it will take some time to load. Then open Replace (Ctrl-h) and enter these strings

Find what string: **(ws="seh".+)kh(.+<)**  
Replace string: **\1k\2**

Make sure the option **. matches newline** is NOT selected.

What the expression means: Search for **ws="seh"**, then any number of characters on the same line, then a kh, then any number of characters on the same line, then a <. Replace with the contents of the first parenthesis in the Find what string, then a hook k, then the contents of the second set of parentheses.

Save your changes, This modifies the **sena 3.fwdata** file directly.

*Wes Peacock points out that the regular expression solution is limited if there is more than one string that needs changing in the same line. When I tested this, he is right, the expression makes only one change per line. I did find if I kept pressing "Change All" in Notepad++ the expression eventually got multiple instances per line changed. The CC solution does not have this problem.*

Examine results:

After either option, you should have a changed **Sena 3.fwdata** file in place in the Fieldworks project directory. You can start FLEEx, open the project and see the changes. If something has gone wrong, you can go back to how it was by closing FLEEx, finding the copy of the original fwdata file you saved in preparation, and copying it back to the project directory, replacing your changed file.