

Importing Scripture into Paratext Using SILAS - Tiki

Author: LR-Language Software Support

Date: 27-Aug-2014 10:13

URL:

<http://confluence.sil.org.pg/display/lcore/Importing+Scripture+into+Paratext+Using+SILAS+++Tiki>

Table of Contents

1	1. Install Silas	4
2	2. Open the File in Microsoft Word and Start SILAS	5
3	3. Use SILAS	7
	3.1 Try "Save as USFM Scripture"	7
	3.2 Format Styles	7
	3.3 Format Verse Numbers	8
4	4. Save and Edit SFM File	9
5	5. Import into Paratext	10
	5.1 Import Books	10
	5.2 Copy/Paste	10
6	6. Fix Backslash Codes using Regex	11
	6.1 Paratext	11
	6.2 RegExPal	11
	6.3 Use EditPad Lite	11
7	7. Helpful Regex Expressions	12
	7.1 Remove empty \p markers at the end of a chapter (EditPad Lite)	12
	7.2 Add missing \p after \s	12
	7.3 Remove extra \p	12
	7.4 Add \p between \c and \v	12
	7.5 Remove empty \p before \s	12
	7.5.1 Find Chapter/Verse Numbers at the End of a Line	12
8	8. Helpful Tips	13
	8.1 Adding Missing References to Footnotes and Cross-References	13
	8.2 Refresh the Project in Paratext	13
9	9. Window Pop-Ups	14
	9.1 No Language Set (LLD)	14
	9.2 Style names from out-of-date stylesheet	14
	9.3 No Language Set (SetDefPaperSize)	15
	9.4 If your current file is NOT a Scripture portion...	15
	9.5 I can't find the standard USFM stylenames in your document.	16
10	10. Page Contents	17

If you have Scripture that is not in SFM format, you can import it into Paratext using SILAS.

1 1. Install Silas

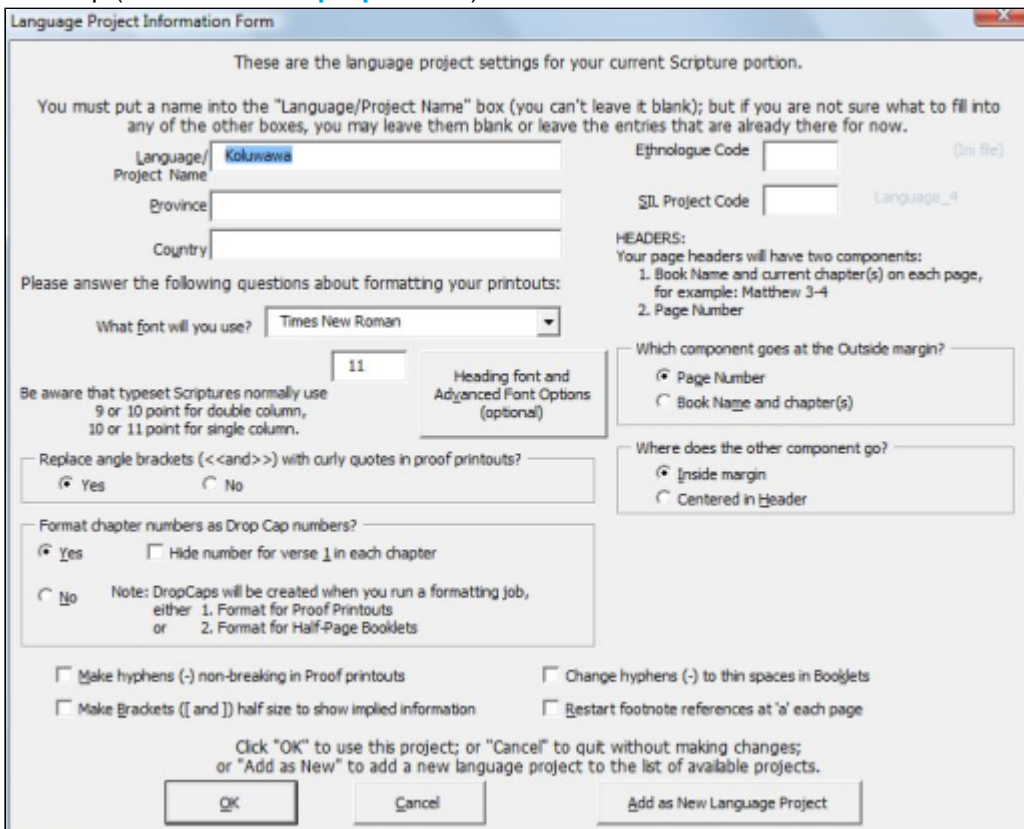
1. Install [Silas](#).

2.2. Open the File in Microsoft Word and Start SILAS

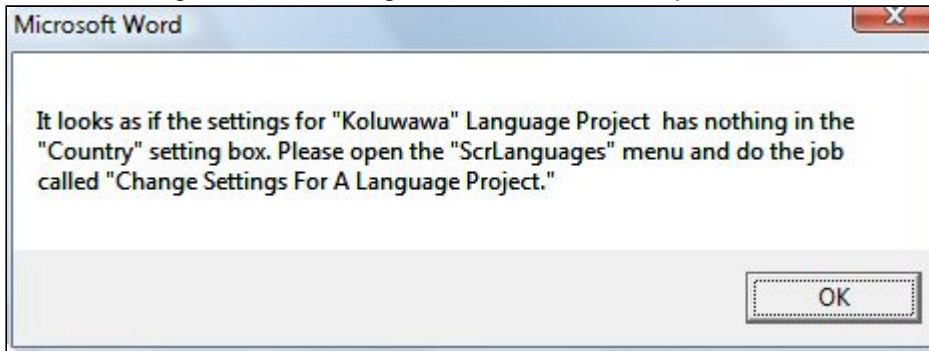
1. Open the Word file you wish to convert
 - a. If you just have *plain text*, open the plain text in Word and **save** it as a **DOC/DOCX**. Silas must have a DOC/DOCX to work with.
2. Click on **Add-Ins > Attach SILAS Formatting Template**



3. If the following window comes up, you can click Cancel unless you want to specify settings for a specific language to import. If you define settings, you may avoid some of the "extra" windows that come up (see [Window Pop-Ups](#) below)



4. You can also ignore the following comment as well. Just press OK to continue.

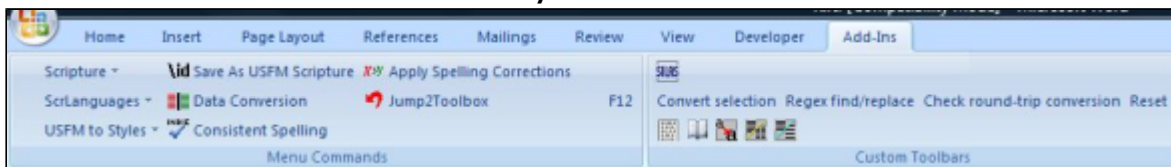


3 3. Use SILAS

From here, you can do a few different things:

3.1 Try "Save as USFM Scripture"

1. Go to the **Add-Ins** tab > **Save as USFM Scripture**



2. If that doesn't work (i.e., if verse numbers do NOT get converted properly try the following), try the next things.

3.2 Format Styles

1. If you have the ability to "Select All Text With Similar Formatting" (I know Office 2003 and up has it) and your verse numbers and/or section headings have specific styles, do the following for each style type (e.g., chapter, verse, section heading)
 - a. **Select** some **text** with the **specified style**.
 - b. Click **Select All Text with Similar Formatting**
 - i. Office 2007/2010: **Right-click > Styles > Select All Text with Similar Formatting**
 - ii. Office 2003: **Right-click > Select All Text with Similar Formatting**
 - c. Open the **Styles window (CTRL + SHIFT + S)**
 - i. Office 2007/2010: This will open the Task Pane
 - ii. Office 2003: This will highlight the Style area in the task bar
 - d. Choose the appropriate style (all start with the corresponding backslash code in Paratext)
 - i. v - Verse Number
 - ii. c - Chapter Number
 - iii. s - Section Heading
 1. Click **ENTER** or the **Apply/Reapply** button
 - e. Repeat for **each style type**.

3.3 Format Verse Numbers

1. Click on **USFM to Styles > Format Verse Numbers**. This will automatically cycle through the document and check all numbers to see if they should be verse numbers. It will ask you to verify if "out-of-sequence" numbers are actual verse numbers or just part of the text.



4 4. Save and Edit SFM File

Once your Word document is as far as you can get it:

1. Go to the *Add-Ins* tab > *Save as USFM Scripture*.
2. Once the "Save as USFM Scripture" has processed, I recommend:
 - a. Adding the \id line to the beginning of the document:
 - i. See [Paratext Book Names and Abbreviations](#).
 - b. Saving the created *TXT file*.
 - c. Closing that TXT file in *Word*
 - d. Opening it in a *good text editor* for further processing
 - i. I highly recommend [EditPad Lite](#) (it's *free*).
3. Sometimes Silas works wonderfully and you don't have to make any changes. Other times, it will put \Chapter instead of \c, or \quote/poetry instead of \q1, etc. I use EditPad Lite to find/replace these changes so they work for Paratext.
 - a. If you know about doing regex replacement, that can be very helpful to process some of the changes that may need to be made (EditPad Lite has great regex functionality). If not, that's ok, it will just take you a little while longer. 😊

5 5. Import into Paratext

After you have all of the changes you want, import it into Paratext. This can be done in 1 of 2 ways.

Note: It can be VERY helpful to create a temporary project where you can fix everything up, then copy the files to the "real" project. This avoids cluttering the real project's history).

5.1 Import Books

1. Project > Import Books

5.2 Copy/Paste

1. Copy all of the text of the TXT file
2. In Paratext (under the project/book you're importing into), uncheck View > By Chapter
3. Choose View > Standard
4. Paste the contents of the TXT file

6 6. Fix Backslash Codes using Regex

MUCH of the work of converting to SFM is done for you by SILAS because it creates virtually all of the backslash codes you need, but there are often invalid backslash codes generated (either remnant from the SILAS export, or from the original data).

You can fix the errors in a number of ways. See [Helpful Regex Expressions](#) below for help with regex.

Be sure to use Mark Point in Project History frequently in case you make a mistake.

6.1 Paratext

1. Run Chapter/Verse Numbers and Markers checks
 - a. This can also help identify errors to fix using RegExPal or EditPad Lite.
2. Do a find/replace for simple code changes
3. View > Standard can be helpful for this because it will highlight errors in backslash codes in red.

6.2 RegExPal

This can help fine-tune your regex and keep a history of what has happened in the case of an error

1. Use the Extract feature before applying the replace to make sure you're finding exactly what you want
2. Use the Replace feature to make the changes you want
3. Close/Re-open the project in Paratext to view changes

6.3 Use EditPad Lite

This is especially useful if there are empty paragraph markers at the end of a chapter

1. Edit in EditPad Lite
2. Re-import.
3. Use Mark Point in Project History because EditPad Lite does NOT keep track of history.

7 Helpful Regex Expressions

7.1 Remove empty \p markers at the end of a chapter (EditPad Lite)

RegExPal can't really be used to find empty \p markers at the end of a chapter because it keeps chapters separate.

You can do the following in EditPad Lite though:

- Find: `\p\r\n\c`
- Replace: `\c`

7.2 Add missing \p after \s

- Find: `(\s[123].*\r\n)(\lv)`
- Replace: `\1\p\r\n\2`

7.3 Remove extra \p

- Find: `(\p *\r\n)+\p`
- Replace: `\p`

7.4 Add \p between \c and \v

- Find: `(\c [0-9]+\r\n)(\lv)`
- Replace: `\1\p\r\n\2`

7.5 Remove empty \p before \s

- Find: `\p *\r\n(\s[123])`
- Replace: `\1`

7.5.1 Find Chapter/Verse Numbers at the End of a Line

- TODO

8 Helpful Tips

8.1 Adding Missing References to Footnotes and Cross-References

Paratext makes it simple to add missing references to footnotes and cross-references

1. Use Checking > Paratext 6 Checks > Publishing > Insert Footnote References.
2. Be sure to click Options and make the following changes
 - a. CV Separator: put a ":" instead of "." if you wish
 - b. Insert \fr: Choose "Yes" to insert references in a footnote (default is "No")
 - c. Insert \xo: Choose "Yes" to insert references in a cross reference
3. **Note:** Before running this, make sure each footnote/cross reference has at least "\f + \ft ..." or "\x - \xt..." or the reference will be inserted after the first word. For example
 - a. *Without \ft*
 - i. Start: **\f This is a footnote\{***
 - ii. End: **\f This \fr 1:2 \fr is a footnote\{***
 - b. *With \ft*
 - i. START: **\f + \ft This is a footnote***
 - ii. END: **\f + \fr 1:2 \fr This is a footnote\{***
4. After running this, you will need to replace \fr* with \ft.

8.2 Refresh the Project in Paratext

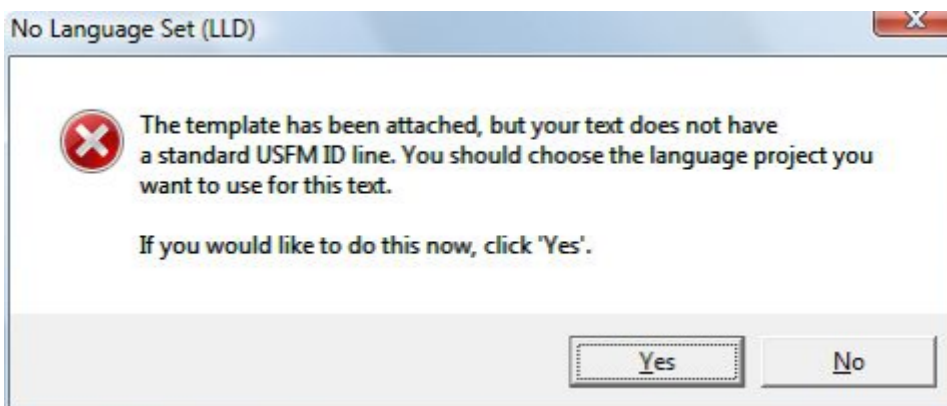
If using RegExPal and Basic Checks in Paratext, you can do the following to refresh the project in Paratext

1. Make changes in RegExPal
2. Close the project in Paratext
3. Press "Rerun" in the checking window: this will re-open the project window as well

9 Window Pop-Ups

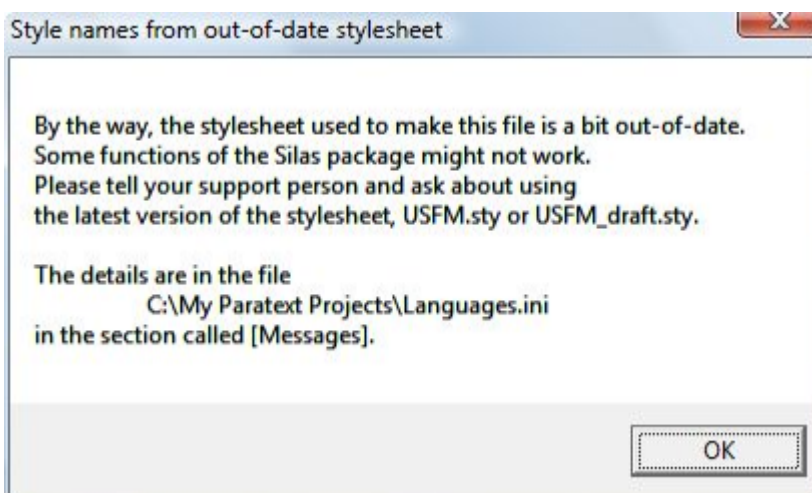
There are number of other possible windows that can come up (and may or may not be consistent) which can be ignored if you want. They are listed below:

9.1 No Language Set (LLD)



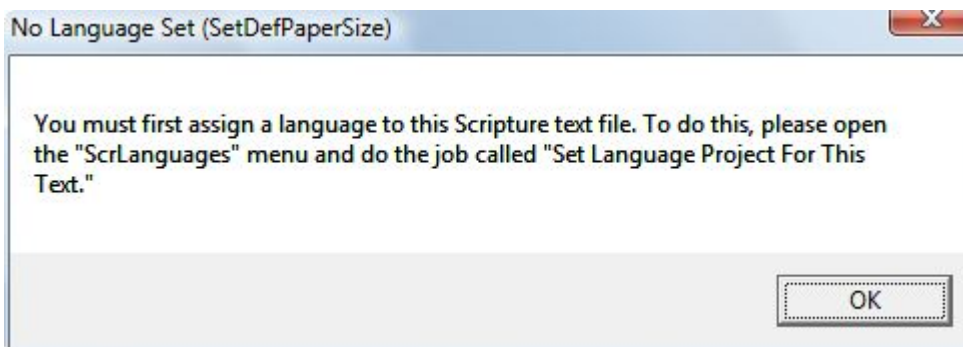
1. This means there is no lid line:
 - a. I find it easier to add it later so you can click "No".
 - b. If you want to define a language, click "Yes".

9.2 Style names from out-of-date stylesheet



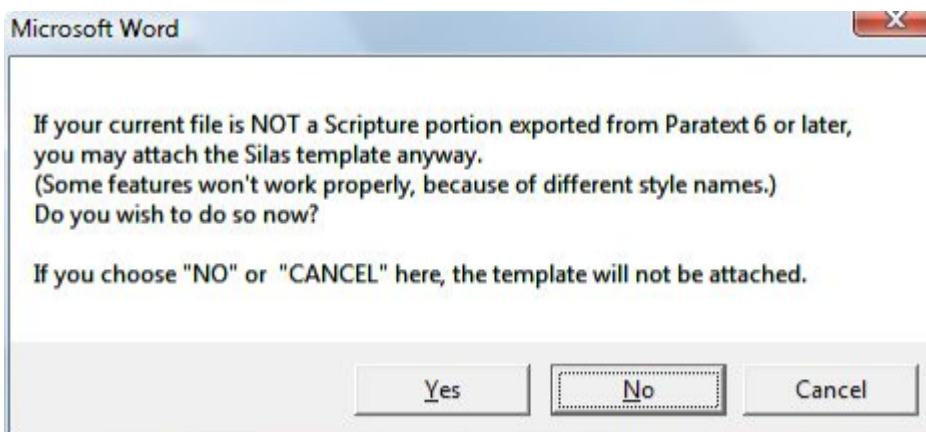
1. Click OK

9.3 No Language Set (SetDefPaperSize)



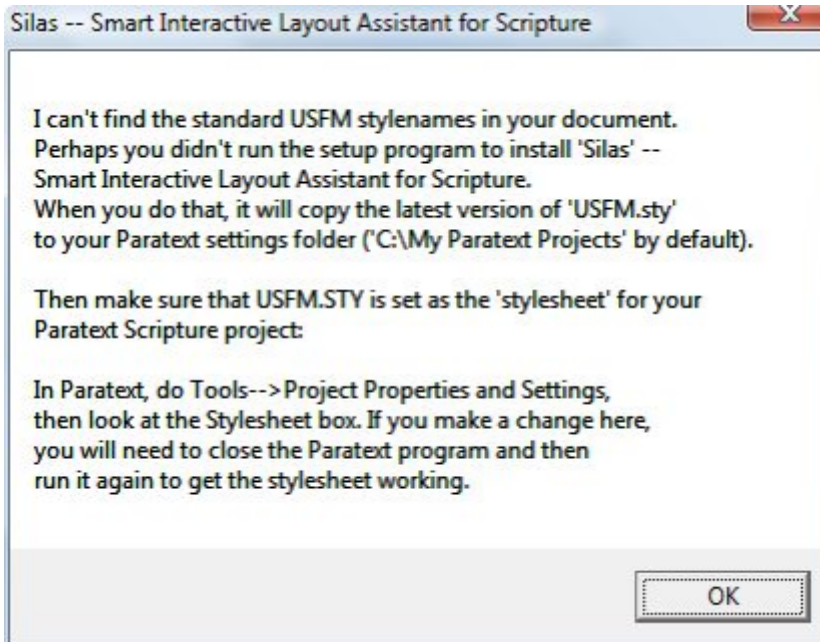
1. This isn't necessary, just click OK

9.4 If your current file is NOT a Scripture portion...



1. Click Yes because you WANT to attach the template anyway.

9.5 I can't find the standard USFM stylenames in your document.



1. Click OK and see if it works.

10 Page Contents

- 1. Install Silas
- 2. Open the File in Microsoft Word and Start SILAS
- 3. Use SILAS
 - Try "Save as USFM Scripture"
 - Format Styles
 - Format Verse Numbers
- 4. Save and Edit SFM File
- 5. Import into Paratext
 - Import Books
 - Copy/Paste
- 6. Fix Backslash Codes using Regex
 - Paratext
 - RegExPal
 - Use EditPad Lite
- Helpful Regex Expressions
 - Remove empty \p markers at the end of a chapter (EditPad Lite)
 - Add missing \p after \s
 - Remove extra \p
 - Add \p between \c and \v
 - Remove empty \p before \s
 - - Find Chapter/Verse Numbers at the End of a Line
- Helpful Tips
 - Adding Missing References to Footnotes and Cross-References
 - Refresh the Project in Paratext
- Window Pop-Ups
 - No Language Set (LLD)
 - Style names from out-of-date stylesheet
 - No Language Set (SetDefPaperSize)
 - If your current file is NOT a Scripture portion...
 - I can't find the standard USFM stylenames in your document.