Help for Linguistic Tools OpenOffice add-on

Overview

The Linguistic Tools add-on is designed to import data in various XML file formats. These files can be exported from Phonology Assistant, Toolbox, and FieldWorks. After importing the data, you can insert language examples into the writeup by typing in the reference number.
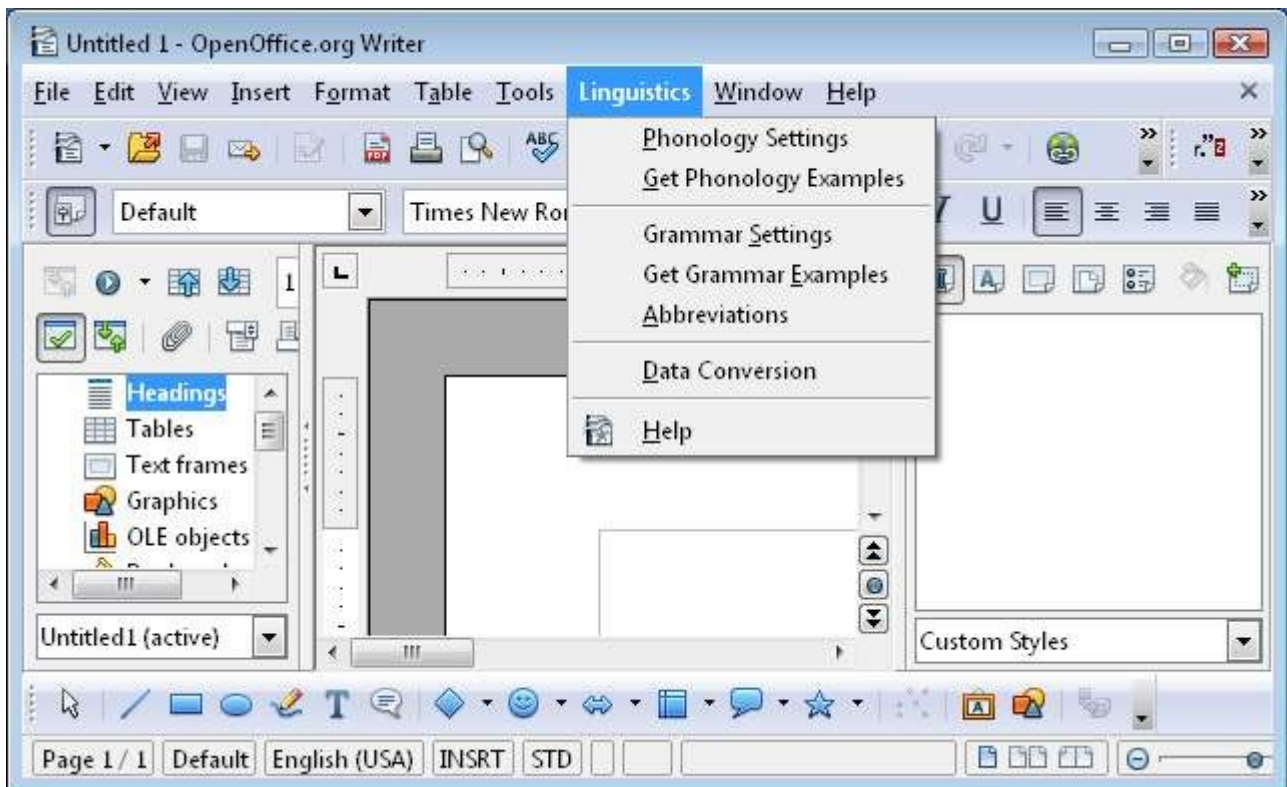
1. **Phonology Settings / Get Phonology Examples**

   Load phonology data from Phonology Assistant or other software and insert examples.

2. **Grammar Settings / Get Grammar Examples**

   Load interlinear data from Toolbox or Flex and insert examples.

3. **Abbreviations**

   Keep track of grammatical abbreviations used in a writeup.

4. Data Conversion

Run SIL Converters to make script changes based on style or font.

In this help document there are also tips for setting up OOo documents. Especially recommended are the tools for maintaining LSA bibliographies and citations.

File formats

To work on the document, please save it in OpenOffice Writer's normal format, which is .odt (Open Document Text). Otherwise settings for the Linguistic Tools will be lost, and other document formatting may be lost as well.

MS Word 2007 SP2 and higher should be able to read .odt files. It is also possible to save a copy in .doc format for someone else to read. However with either of these approaches there are usually some formatting problems. The best method of saving a writeup for someone else to read is to export as PDF (File -> Export as PDF).

If possible it is best to start with an OpenOffice document rather than importing from MS Word. Importing may be your best option, but if so, be warned that there will be more problems that way.

Conventions

Edit -> Copy	This means "In the Edit menu, click on Copy."
OOo	This is an abbreviation for the OpenOffice.org suite of applications. Unofficially it is often referred to as OpenOffice.

Installation

The add-on is installed from a file called LinguisticTools.oxt. When you double-click on this file, the OpenOffice Extensions Manager will open and install it. After it finishes, open a Writer document. If the installation was successful, you will see a new menu to the right of the Tools menu called "Linguistics" as shown in Overview above.
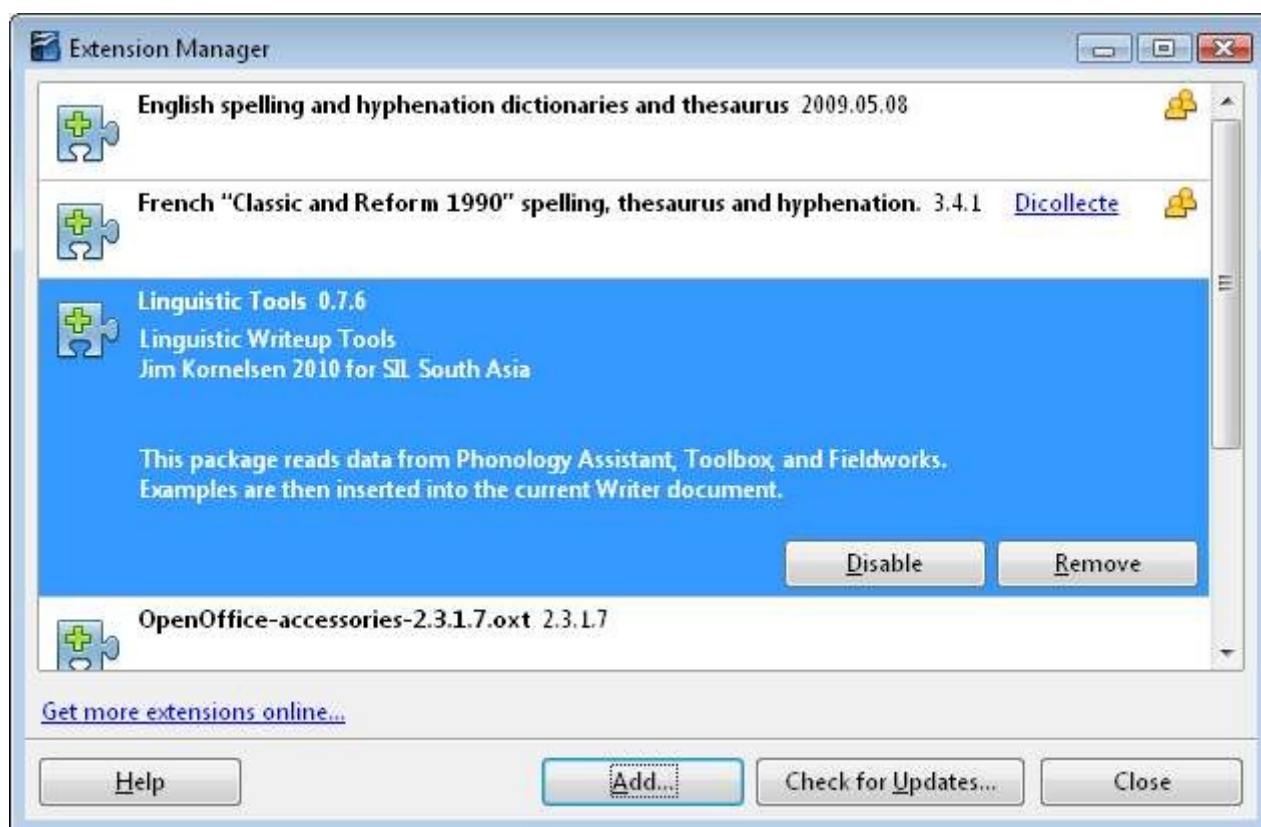
If the extension is not able to install correctly, close all open windows of OpenOffice.org and then try again. You may also try rebooting if there are problems. To use Data Conversion, see additional installation instructions below.

When installing a new version of OOo, in order for this extension to work, the Python-UNO bridge must be selected. This is under Optional components and is selected by default.

If there is an older version of the add-on installed, the Extension Manager will first remove the old version. You may see an error when removing that says "Addons.xcu does not exist." Just ignore this error and press ok. To prevent this error, you can find out more information in "Reset user's application profile" in the document called "Troubleshooting OpenOffice.org" on the internet.

Sometimes removing an older version may give other error messages such as failing to close a bridge. If there are error messages like this, then close all open windows of OOo. Then try removing the older version again. You may also need to reboot in order for this problem to go away.

Occasionally it may be necessary to manually delete old extension files. The files are located in the OOo user directory. To find this directory, go to Tools -> Options. Under OpenOffice.org, select Paths. Most of the paths listed point to the user directory. Typically it looks like C:\Users\ {username}\AppData\Roaming\OpenOffice.org\3\user\. In the user directory, go to uno_packages\cache\uno_packages. There is a separate folder for each installed extension. Look carefully and delete the old folder along with its .tmp file. Then close and open OOo and reinstall the extension using Extension Manager.

OpenOffice.org version 3.0 or later is required. Some features may only work correctly under Windows, because OOoLT has not yet been well tested on other platforms. For Data Conversion, Windows is required, because SIL Converters requires Windows.

Add-on instructions

Phonology

We will start by describing how to use OOoLT from Phonology Assistant. In Data Corpus view, make sure that the "Reference" column is not empty. Also Phonetic, Phonemic, and Gloss columns should be used. In the File menu, choose Export -> To Phonology Assistant XML. Save the file someplace where you can find it again, for example in My Documents. This will create a file ending in *.paxml. Then close Phonology Assitant.

If the data is coming from Flex, Phonology Assistant will not recognize the Reference column. Instead, you can export directly from Flex to OOoLT. Use the lexical entry field for the phonemic, and pronunciation for phonetic. Alternatively it is possible to use citation form for phonemic and lexical entry for phonetic. Export as LIFT format. You will probably want to save it in a folder with an appropriate name such as "PhonData," because several files will be created for LIFT format.

It is also possible to export directly from Toolbox. The following fields are required: \pht (Phonetic), \phm (Phonemic), \ge (Gloss English), \ref (Reference Number). Export as XML. To do this, go to File -> Export, and add an XML process if it is not there yet. Then press OK to export.

After a data file is exported, open OOo. In the Linguistics menu, choose Phonology Settings. Click on the Browse button to open the file selection dialog[1]. Browse to the data file that you exported (for example the file ending in *.paxml from Phonology Assistant) and press Open. Words will be shown between slashes, brackets and quotes unless the box is unchecked. Also you can choose whether the phonetic or phonemic representation comes first.

Click Ok to close Phonology Settings. Next, you can type reference numbers that refer to the data, for example #0016 or #COCO.01. Then go to Linguistics -> Get Phonology Examples. Check the box "Search from beginning of document" and click Find Next Reference. Now click Replace with Example. It should load the example into the document. If you have typed several reference numbers, then again press Replace with Example.

Another method is to type in a ref number in the Get Phonology Examples dialog window and press the "Insert This Example" button. Some people find this method easier.

---

1   There is a bug in OpenOffice version 3.0 on Windows Vista that prevents the file selection dialog from opening. A workaround is to go to Tools->Options, select OpenOffice.org->General, and enable the checkbox "use OpenOffice.org dialog for Open/Save." However the best way to fix this problem is to upgrade to OpenOffice 3.1 or higher.

After you have inserted an example, you can insert a cross-reference to the example that will update automatically. Go to Insert -> Cross-Reference and select type as AutoNr. Select the number in the list that was just inserted. Specify to insert reference to Reference. Then click the Insert button, and click Close.

To change font size or color, use the Styles and Formatting pane. If you do not see the Styles and Formatting pane, go to Format -> Styles and Formatting. In the pane, change the box at the bottom to view Custom Styles. Select the Character Styles icon at the top of the pane (the second icon from the left). Right-click on the style, for example Phonemic, and select Modify.

To change tabs, you should modify the Lex Example style. If you instead change tabs directly on the paragraph, then it will only change one example at a time. To modify the style, at the top of the Styles and Formatting pane, select the Paragraph Styles icon (the first icon). Right-click the Lex Example style and go to the tab named "Tabs." In the box labeled "Position," you will see the current positions of the tabs. To move one of them, create a new tab by entering the new position and then pressing "New." Then select one of the old tabs and press the button to delete it.

To update examples, go to Linguistics -> Get Phonology Examples and specify to Update existing examples. Press Find Next, which will find the reference number of an example that you have put into your document earlier. When a reference number is found, press Update Example to get the new data.

If you are confident that updating examples is working well, then you can try pressing Update All, which will attempt to find and update every example in the document. Before doing this, it is a good idea to save a copy of the document in case there are problems. You may have to wait some time for this process to finish.

Grammar

For grammar, we will start by describing how to get data from Toolbox. You can export an XML file from an interlinear text in Toolbox. Interlinear text fields must be set up in a hierarchy. (By the way, a hierarchy helps with interlinearizing as well). If the hierarchy is not set up properly, then the data will not be imported correctly into OpenOffice. To see the hierarchy, go to View -> Marker Hierarchy, and View -> Both Markers and Names. The fields should look like this:

\id Identification
. \ref Reference Number
. . \tor Orthographic Text
. . \tx Text
. . . \mb Morpheme Breaks
. . . . \mor Orthographic Morpheme
. . . . \ge Gloss English
. . . . \ps Part of Speech
. . \ft Free Translation

To change the hierarchy, go to Database -> Properties. Modify markers so that they are under other markers. For example, select \ge and press Modify. Change "Under what in the Hierarchy" to "mb Morpheme Breaks" and press OK.

If you want to use markers that are different from what is shown above, it is possible to set up OOoLT to use different markers, as long as the hierarchy is set up properly. Modify settings as described in Advanced settings. If the gloss marker is "gl" for example, then change the value of LTg_SFMarker_Gloss to "gl" instead of "ge."

You can also export interlinear texts as XML files from Flex. Each text will be in a separate XML file. Reference numbers are automatically numbered starting from 1.1 in each file, which makes it difficult to refer to a specific example. For a solution to this problem, specify additional prefixes as described below.

Now in OOo, go to Linguistics -> Grammar Settings.

If you uncheck the Morphemes in separate columns, then morphemes will be separated by dashes. You can uncheck Text or Morphemes if you do not want that line included.

If you are using Flex and have several texts, you can specify a Ref Number prefix for each file. For example, the prefix for the first file could be "S1-" (short for "story 1"). Then you can refer to the first example (number 1.1) by typing "#S1-1.1". If you are using Toolbox or you have only one FieldWorks text, then there is probably no need to set a prefix.

There are two methods for output: frames and tables. Frames are more flexible for changing font size or removing words. However some people prefer tables in order to select rows and columns more easily If you want to save a copy of the document in .doc format, you may find that it is necessary to use tables.

By default example numbers are inserted in a separate column. This makes it line up nicely. If you need to use high example numbers (over 1000 for example), you may need to increase the column width. When you change the column width, Linguistic Tools will attempt to resize all examples already in the document.

In order for examples to be in a separate column, an outer table is created. This is normally recommended. However if you need to put several examples inside of another table, this may not be what you want. To avoid this, uncheck Outer table. Example numbers will still be added, but they will not be lined up as well. A drawback of unchecking Outer table is that it will not be possible to update these examples automatically. More information on updating examples is given below.

If you do not want example numbers at all, then uncheck Numbering.

Many people who are fairly competent with computers like to add orthographic data to their writeups. In OOoLT, there are two basic kinds: Word (Text) and Morpheme data. In FieldWorks, these lines are just an additional writing system, and actually it doesn't matter which one is which, because OOoLT will just keep the order from Flex.[2] In Toolbox, use \tor and \mor, and make sure they are at the correct position in the hierarchy. It is not currently possible to change the order in Toolbox. Often in Flex and Toolbox, orthographic lines are generated by lookup or conversion processes. It is beyond the scope of this document to give instructions for these processes.

To insert examples, go to Linguistics -> Get Grammar Examples. This is similar to the method of getting phonology examples described above.

To change font size or color, go to the Styles and Formatting pane and select Custom Styles in the box at the bottom. Select the Paragraph Styles icon at the top. Right-click on the style, for example "Interlin Base," and select Modify. If you are using frames, then the frames will automatically adjust as you change the font sizes.

If you are using tables, it is ok to make the font smaller, but making the font bigger will often cause the morphemes and words to be too big to fit, and then it will not look good. In this case it is probably necessary to update all of the examples after making the font size bigger.

If the spacing of the table doesn't look right, please check the "Table Contents" paragraph style. Right click on the style and select Modify. In the Indents & Spacing tab, everything should be set to 0.00, and line spacing should be Single. Often documents from MS Word have this style set incorrectly.

---

2 At least I think this is how it will work, although I haven't verified this yet.

Updating grammar examples

If Outer table is checked in Grammar Settings, then it should be possible to update examples using the Linguistics menu. Go to Linguistics -> Get Grammar Examples, and specify to Update existing examples. Press Find Next, which will find the reference number of an example that you have put into your document earlier. When a reference number is found, press Update Example to get the new data. This will do a three things:

1. Insert an updated example.
2. Delete the old example (but it will keep the example number field, so cross references should not be affected).
3. Put a copy of the old and new examples in a blank document for checking.

After you have updated a number of examples, you should go to the blank document that gets opened automatically. In this document you can review the changes, and if something is not quite right, you can click the button to go back to the main document and fix the problem. For example, you may notice that one of the words should be formatted in Bold, in which case you need to do this again in the updated example. When finished checking for problems, close the blank document without saving. Each time you go to Linguistics -> Get Grammar Examples and update an example, a new blank document is opened, so make sure you close all of them.

It is recommended to use the blank document for checking. However if you really don't want the blank document to be created, it is possible to disable this. To do so, follow the instructions in Advanced settings and set LTg_ComparisonDoc to 0.

If you are confident that updating examples is working, then you can try using the Update All button. However this may cause problems. So before updating all grammar examples, make sure to save a backup copy of your document. Then click Update All. You may be able to see when each example is updated, but often OpenOffice will stop responding because it is busy updating all of the examples. When this happens, please wait for about 10 minutes or so to let it finish. If after a long time it is still not responding, there may have been some problem. In that case you may need to force OpenOffice to close.

If the "Update examples" feature does not do what you need, here is a more advanced method:

1. Set the Interlin Reference Number style to a font that is different from the rest of the document, for example Tahoma or Arial Black.
2. Go to Edit -> Find and Replace. Click the button to show More Options. Click on the

Format... button. Set the font from step 1 and press Ok. Then check Including Styles. (If you do not see the "Including Styles" option, then close the Find dialog, move the cursor outside of any tables, and then try again.)

3. Press the Find All button. Now all of the reference numbers should be selected.. Press Close, and Copy the reference numbers (ctrl + c). Open a new Writer document and then paste them (ctrl + v).

4. Now we need to add # in front of each number. To do this, go to Edit -> Find and Replace. Click more options, and click No Format (to clear the search from step 2). Check the Regular Expressions box. In the Search For box, type one space followed by {1,4} (open curly brace + one + comma + four + close curly brace). In the Replace With box, type \n# (forward slash + n + number sign). Press Replace All.

5. In the new Writer document, go to Linguistics -> Grammar Settings and specify the file and font settings just as in the main document. You may also want to uncheck Numbering, which means you will need to copy and paste the old example numbers. The advantage of this is that any cross-referenced example numbers will be kept.

6. Now go to Linguistics -> Get Grammar Examples. First press Find Next Reference and then Replace with Example. If it is working correctly, then click Replace All. As with updating all examples, OpenOffice may stop responding for awhile, in which case you probably need to wait until it is finished.

7. Now cut and paste the examples into the writeup. You will probably want to check each example one at a time, because there may have been changes.

Grammatical Abbreviations

One of the things required for a grammar writeup is maintaining a list of abbreviations. A list is normally kept either at the beginning or the end of the document that tells what the abbreviations mean. However this list is difficult to maintain, and it is easy to be inconsistent with labels. To help with these things, use the Abbreviations tool.

When you first open Abbreviations, it will begin with a default list taken from the Leipzig Parsing Rules. Please go through this list first and make some corrections. Don't worry, if the abbreviation is not used in your document then it will not show up later. To correct an item in the list, select it, change the abbreviation or full name, and press update. If it is not in the document but you want it to be included in the list anyway, check Output even if not found and press update.

If the abbreviations are all CAPS and this is not what you want, there is a button to change the list to either lowercase or just the first letter capitalized. Press the button to change the list from CAPS to lowercase, and press it again to change from lowercase to first letter capitalized.

To add new abbreviations, type in the information and press the Add New button. You can also use the Delete button to remove abbreviations, but there is no need to delete all of the unused abbreviations because they will not show up.

Now click "Scan for occurences." This will search through the document and count how many times each abbreviation is used. If the abbreviation is not found, then it will not be included when the list is inserted.

There are probably more abbreviations in your document that you haven't found yet. That is what the upper right-hand section of this window is for. Select a paragraph style to search in, probably either Interlin Gloss or Interlin POS. If your language is head final (typical in South Asia), then leave suffixes selected. Or you can specify Any which will give more results. You can reduce the maximum length to 4 or 3 to look for short abbreviations. Also if the abbreviations are all upper case (for example GEN for genetive), then check upper case.

Now click Search. If an abbreviation is found that needs to be added to the list, click Yes, otherwise click No. Keep doing this until all of the abbreviations are found that you need. After this is finished, go through the list and add full names for the abbreviations. Make any other corrections that are needed.

When the list is finished, click "Insert List" to insert the list into the document at the present location. The list uses a paragraph style that you can use to format it.

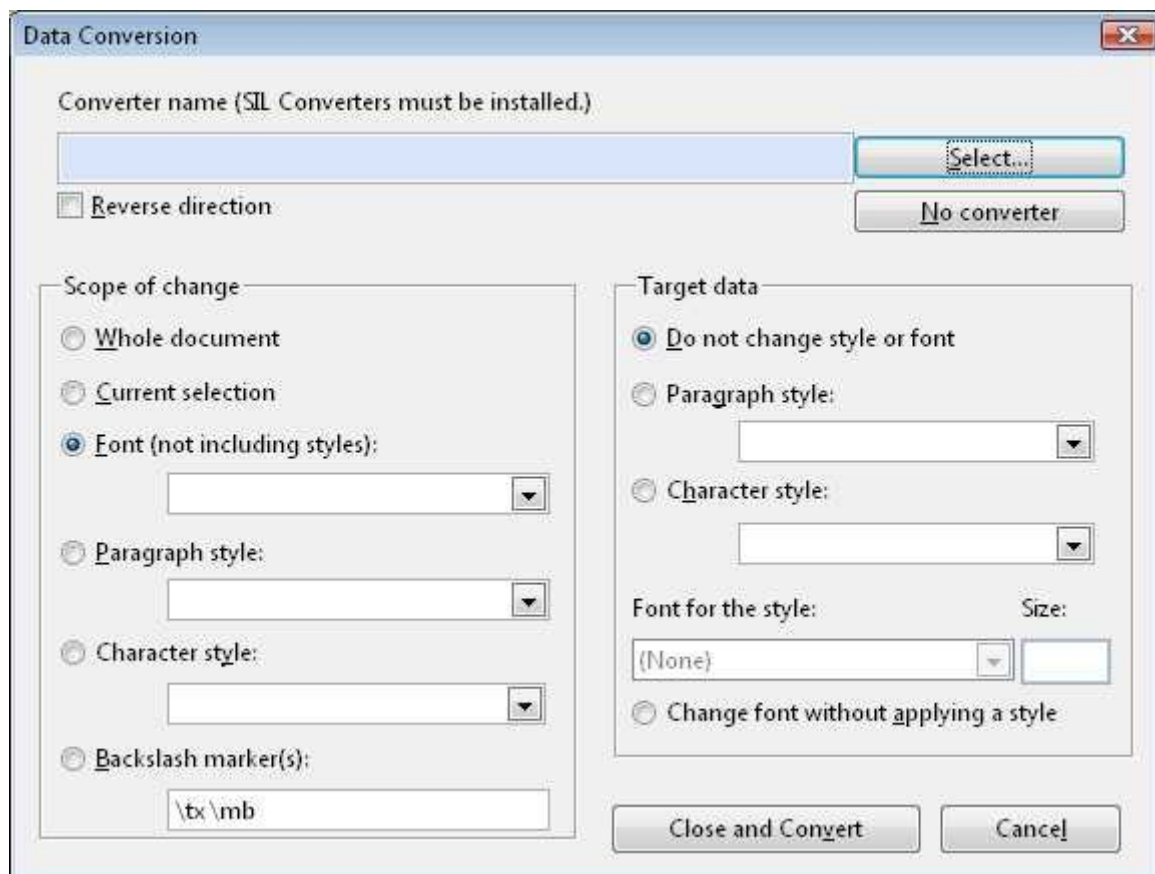If you press "Close" then any changes to the list of abbreviations will be saved.

Data Conversion

In order to use Data Conversion, please download and install SIL Converters.

The first time you try using Data Conversion after installing SIL Converters, it may cause a crash. If this happens, please reboot your PC. It should not happen after that. (Technical note: This is because part of the Data Conversion tool is implemented in a DLL file. OOo seems to have trouble loading the DLL without rebooting.)

In the documentation for the MS Word Data Conversion tool, it says that sometimes the reference to the DLL must be updated. However the OOo tool loads the DLL dynamically, so there is no reference that needs to be updated.

After it is installed correctly, go to Linguistics -> Data Conversion.

Clicking Select should bring up the SIL Converters window, where you can add new converters or make changes. If you want to convert from the right encoding to the left one instead of the normal left to right, then check Reverse direction. After selecting the converter you want to use, press Ok to close the SIL Converters window.

If you want to use the Data Conversion tool to set a new paragraph style but you don't want to do conversion, then click the No Converter button. Keep in mind that OOo's Find and Replace menu has a number of features to make powerful changes to styles and formatting, so that is often a better approach than using the No Converter button.

Specify the scope of change that you want to do the conversion on, either an entire selection or a specific font or style[3]. It is also possible to open an SFM file and convert by backslash markers.

To convert to a new style, first create the style in the Styles and Formatting pane. Then go to Linguistics -> Data Conversion, click the appropriate style radio button for the target, and select the style you created. Alternatively in the Data Conversion window, you can type in the name of a new style and select its font, and OOoLT will create the style.

Often it is better to convert to a style in case changes need to be made to the font or size later. However if you don't want this, select "Change font without applying a style." Then the font will be changed but no style will be created.

One useful approach with the Data Conversion tool is to first use OOo's Find & Replace menu to set a style. You can search by text, color, size, or almost anything else. To do this, go to Edit -> Find & Replace. Click More Options, and click the Format button. Specify the color, size, or font that you want, then press OK. Now press the Find All button, then press close. The results should still be highlighted. You can set the highlighted results to a particular style. Now use the Data Conversion tool to make changes to that style.

After making changes to the settings, click Close and Convert.

---

3 Searching for paragraph styles should work well. Searching for character styles is not a built-in part of OpenOffice, so it has some limitations. For example it will not find character styles located inside Frames.

Advanced settings

If you are careful, you can make advanced changes to the settings of the Linguistic Tools add-on. To do this, first set basic settings if you have not yet done so, for example Linguistcs -> Phonology Settings. Next, go to Insert -> Fields -> Other, or press Ctrl + F2. Go to the Variables tab and select User Field. All of the variables beginning with "LT" are for the Linguistic Tools add-on.

One example is that you could change which style is used for interlinear glosses. Select the LTg_GlossStyle variable. The normal value is "Interlin Freeform Gloss." However you could use a different style instead, for example Gloss1. Change the value to Gloss1 and press the green checkmark to apply. Now the style Gloss1 will be used instead.

Formatting writeups

Here are some suggestions for using OOo to format linguistic writeups.

Navigator

Many people want to easily see the headings in their document. To do this, go to View -> Navigator. It is best to "dock" the Navigator window so that it is attached to the main window. This can be done by dragging with the mouse, although it is somewhat difficult. Click in the title bar of the Navigator window (the blue part), and drag it around the left edge of the main window. When you see an outline appear, then it will dock when you release the mouse button.

Cover page

It is probably best to use the Styles and Formatting pane also. To do this, go to Format -> Styles and Formatting. This will show a window pane on the right side where you can view all of the styles.

In the Styles and Formatting pane, click on the Page Styles icon at the top. Now right-click on the white area in the pane, and select "New..." This will bring up the Page Style dialog window. Change the name to "My Cover Page." Click Ok to create the new style.

Also create two other page styles called My Table of Contents and My Document Body. Then modify the My Cover Page style so that My Table of Contents is the Next Style, and modify the My Table of Contents style so that My Document Body is the next style.

Go to the first page and double-click "My Cover Page" to apply the style.

To center the title on the cover page, go to Insert -> Frame. Set Anchor to Page, and choose Center for both Horizontal and Vertical position. Then press Ok. Now you can type the title in this frame.

Then insert a page break (Insert -> Manual Break), selecting the style as My Table of Contents.

Page numbering

Go to Format -> Paragraph and click on the Text Flow tab. Under the Breaks section, check Insert and With Page Style, and set Page number to 1, and Style to My Document Body. The click Ok.

Now go to Insert -> Footer -> All. Click to move the cursor in the footer. Then go to Insert -> Fields -> Page Number. If you set up the page style correctly, this should show the page number as 1.

To show the total number of pages (for example "Page 1 of 17"), click in the footer and press F2 to bring up the formula bar. Enter the formula "=page – 3" (without quotes). You will need to change this number 3 to the number of pages contained in your document before the body begins.

To see the new page numbers, go to Tools -> Update -> Fields (or press F9). Or, to get page numbers to update, File -> Page Preview.

You may also want to add page numbers that look like "iii" to the table of contents pages. To do this, insert a footer and page number field. Then double-click just before the page number field. This brings up the Edit Fields dialog. Select the Roman (i ii iii) format and click ok.

Warning: The offset field in the Edit Fields dialog provides a quick way to change the page numbers to start at 1. However, do not use this feature. It is better to use page styles.

You may notice that inserting a page break that creates two odd numbered pages in a row, for example iii followed by 1, causes an extra blank page to be printed. If you have this problem, go to Tools -> Options, and under OpenOffice.org Writer, select Print. Uncheck "Print automatically inserted blank pages." Later, if you want to Export as PDF, you will want to uncheck "Print extra blank pages" in the PDF options.

Table of contents

On the table of contents page, go to Insert -> Indexes and Tables -> Indexes and Tables. To change how the Table of Contents looks, go to the Entries tab. Here are some suggestions:
- Click in white space before E, press Hyperlink button.
- Click in white space after E, press Hyperlink button.

- Click in white space after E#, use space bar to make one or two spaces.
- Click on LE
- Select a style of just Arial, no underline. (I made my own style for this purpose).
- Click All button
- If you need to delete a button, you can click on it and then press the Del key.

Also in the Styles tab, you may want to make some changes:

- 1.5 line spacing for Headings 1 and 2
- Add spacing before Heading 1

Outline numbering

For linguistic writeups, many people use outline numbering for section headings, for example section 2.1. A good way to do this is to go to Tools -> Outline Numbering.

In the Outline Numbering dialog box, first select level "1-10." This modifies all levels. Select "1,2,3" for the number format. Change Show sublevels to 10. In the Separator, After field, type a period (.) followed by a space ( ).

Next, select level 5. Set number to None, and show sublevels to 1. Leave the Separator, After field blank. Click Ok to make changes.

I found it helpful to use heading 5 for acknowledgements and bibligraphy. In this case, you may want to edit the heading 5 style. In the Styles and Formatting pane, click the Paragraph Styles icon at the top. Right-click on Heading 5 and choose Modify. In the Indents and Spacing tab, set Before text to 0. In the Font tab, set size to 14 (or whatever you choose).

LSA bibliographies

If you have Firefox, then I recommend using a plugin called Zotero for citations and bibliographies. It has the LSA stylesheet built in. Go to the web site http://www.zotero.org/ and download the latest stable version, and install it as a plugin in Firefox.

Then go to Cite from within Word and OpenOffice and scroll down on that page and click on Installation. Download and install the OOo extension.

If you want to use LSA format, you can download a stylesheet. Click on the link from the Zotero home page to use thousands of [bibliographic styles](). Scroll down to Language [LSA] and click the Install link.

To use Zotero, first add bibliography entries in Firefox. Then in OOo, go to View -> Toolbars, and show the Zotero toolbar. You can use the toolbar buttons to configure Zotero and insert citations and a bibliography.

If there is an error connecting to Firefox, then make sure Firefox is open. If the computer is not connected to the internet, then in the File menu of Firefox, uncheck Work Offline.

Other tips

To create a new style, first select some text and format it the way you want it, for example Bold size 14. Then in the Styles pane, click the "New Style from Selection" button.

To set some text to a different style, select the text, then double-click on the style. If you are mostly working with custom styles, then you may want to select "Custom Styles" in the box at the bottom of the Styles and Formatting pane.

To insert a cross reference to a heading, go to Insert -> Cross Reference. Select type as Headings. Then select the heading that you want to reference to. Set "Insert reference to" as "Number."

To delete a page break, it may not be enough to press Backspace. If Backspace does not work, go to Format -> Paragraph, and under the Text Flow tab uncheck the Insert box. Paragraphs, tables and other objects can force a paragraph break in this way.

To change page settings such as landscape, go to Format -> Page (not in the File menu like some other applications).

Technical notes

If you are interested in seeing how the add-on is made and perhaps making changes, then simply unzip the .oxt file. You may need to change the extension to .zip first. Most of the code is in Python, under the pythonpath directory. You can make changes, zip it back up, and then reinstall it in OOo extensions manager.

Contact

For questions or feedback, please contact <jim_kornelsen@sil.org>