

# Publishing FLEx Dictionaries Using Microsoft Word via Pathway

Ken Zook

January 9, 2023

## Contents

Publishing FLEx Dictionaries Using Microsoft Word via Pathway.....	1
1 Introduction.....	2
2 Exporting a dictionary to Word via Pathway.....	2
2.1 Preparation in FLEx prior to export.....	2
2.2 Exporting a dictionary from FLEx via Pathway to Word.....	3
2.3 Additional processing of the .doc file before saving in Word.....	5
2.3.1 Enable div classes to import into Paragraph styles.....	6
2.3.2 Insert forced breaks from style before/after text.....	7
2.3.3 Replace utf-8 “language”.....	7
2.4 Preparing a Word dictionary for publication.....	7
2.4.1 Page layout.....	7
2.4.2 Removing hyperlinks.....	8
2.4.3 Embedding pictures in the .docx file.....	8
2.4.4 Adjusting styles.....	8
2.4.5 Add page headers with guidewords.....	9
2.4.6 Convert letter headers to single column.....	11
2.4.7 Editing the internal Word docx file.....	11
2.4.8 Dictionary reversal indices.....	12
2.4.9 Dictionary front matter and back matter.....	12
3 Useful features in Word.....	12
3.1 Basics of Styles in Word.....	12
3.2 Preventing styles from changing: dynamic updating.....	13
3.3 Advanced Find and Replace.....	13
3.3.1 Finding styles.....	14
3.3.2 Replacing styles.....	14
3.3.3 Special codes.....	14
3.3.4 Wild cards.....	15
3.3.5 Adding text before or after a style.....	17
3.4 Saving pictures in the Word document.....	17
3.5 Removing hyperlinks in Word.....	18
3.6 Removing unused styles in Word.....	18
3.7 Running macros in Word.....	19
3.8 Editing the internal Word docx file.....	19
4 Bidirectional dictionaries in Word.....	19
4.1 Enabling right-to-left publications in Word.....	20
4.2 Basic RTL layout in FLEx and Word.....	21
4.2.1 Settings in FLEx.....	21
4.2.2 Settings in Word.....	21
4.3 Bidirectional algorithm.....	21

4.4	Formatting bidirectional entries in FLEEx .....	23
-----	---	----

## 1 Introduction

Microsoft Word is not the ideal publishing software, but it does provide enough capability to provide a reasonably good publication of a dictionary from FLEEx. In Word you can add dictionary introductory sections, reversal indexes, and appendices. It also allows the user to make changes to the FLEEx dictionary that might not be possible to do with the FLEEx configuration options. Once a dictionary is in Microsoft Word, it can easily be exported to PDF. For more sophisticated publications, one possibility is importing the Word document into Adobe InDesign which provides more powerful publication capabilities.

As of FLEEx 9.1.24, FLEEx does not provide a direct export to Microsoft Word. However, using Pathway 1.18.1 or later (<https://software.sil.org/pathway/download/>) there is a path to import a FLEEx xhtml dictionary into Microsoft Word. There are still a number of challenges involved in this process. Some of these challenges may require a skilled user that can modify internal files in Word. This document describes ways to work around most of the challenges, which can get very technical in some cases. It describes some useful features in Word for working on dictionaries. It also describes special features related to bidirectional dictionaries in Word. This document is based on Microsoft Word which comes with Microsoft Office 2019. Exporting one or more reversal indexes needs to be done as separate exports and then the results combined in your final Word document.

## 2 Exporting a dictionary to Word via Pathway

### 2.1 Preparation in FLEEx prior to export

Prior to exporting a dictionary from FLEEx, you need to set everything up to where the Dictionary view in FLEEx shows the entries the way you want them printed. This involves the following steps.

- Select the entries, pronunciations, senses, examples, and pictures to include in the published dictionary. If needed, you can set up a unique publication for this purpose by adding a new publication to the Publications list in the Lists area. The default publication that comes with FLEEx is “Main Dictionary”. FLEEx provides the following fields used to determine if these elements will be included in a particular publication.

Publish Entry In

Publish Pronunciation In

Publish Sense In

Publish Example In

Publish Picture In

If a publication is listed in these fields, it means the element will be included in that publication. These fields can be changed via Bulk Edit Entries.

- The appearance of variants, complex forms (often subentries), and minor entries can also be configured on each entry using these fields

Show As Headword In

Show Subentry under

Show Minor Entry

Subentries

Referenced Complex Forms

- It's possible to set up Filters on one or more columns to limit the entries that occur in the publication. For testing purposes, it's often helpful to set a filter to just a few entries to save time in making test exports to Word. Normally you will want to disable all filters prior to the final export so that you will get all of the desired entries in the dictionary.
- Dictionary sorting. Normally you will want to sort on the Headword field for publications. If the sort order is not correct, use Format...Set up Vernacular Writing Systems...Sorting to adjust the collation rules. Also, the collation rules determine which letter headers will be included in the printed dictionary. Also, letter headers will only be included if you are sorting on a lexeme form or citation form field.
- Use Tools...Configure...Dictionary to configure the dictionary to format the dictionary as desired for the publication. This includes selecting the fields you want to display, the order of the fields, the style of the fields, the writing systems for the field, any special text before or after the field, or between elements in a field. There are also options for formatting homograph numbers, sense numbers, whether some elements are in separate paragraphs or not. Associated with this may also require settings in Format...Styles, including creating custom styles as needed. All of the formatting, except styles can also be set for a unique publication by creating a custom views using the Manage Views button in the Configure Dictionary dialog.
- Make sure the correct view is chosen at the top of the Configure Dictionary dialog.

See section 4.4 for special considerations for bidirectional dictionaries

## 2.2 Exporting a dictionary from FLEx via Pathway to Word

The basic process for exporting a dictionary to Word involves exporting the Dictionary view from FLEx to XHTML, running a Pathway batch file to convert the XHTML to a Word .doc file, and then opening that file in Word. Here are the steps.

1. First, make sure FLEx is using the correct sort and filters in Lexicon Edit. Then in the Publication view blue title bar, make sure the correct publication is selected in the left chooser, and the correct view is selected in the right chooser.
2. To export to a XHTML directory, use File...Export...Configured Dictionary Web page (XHTML). When choosing an export location, make sure you create or use an empty folder for the export, type a file name, and click Save. A separate directory is needed because the export may include an "AudioVisual" subdirectory with audio files, and a "pictures" subdirectory for picture files. In addition to the xhtml file, it will also have 2 style sheets with .css extensions. Note, the xhtml and css files will be overwritten by the next step, so if you want to keep the exported contents, make a copy of the directory before the next step.
3. Converting the xhtml export to a Word .doc file involves running a batch file installed with Pathway. The batch file is located at "C:\Program Files (x86)\SIL\Pathway\FLExXhtmlToDoc.bat". You can use this file directory, or you can create a shortcut to the file in a more convenient location. To execute the batch file, drag the .xhtml file you exported and drop it on the shortcut or the original batch file. This will open a couple Cmd windows during the process. It uses a CssSimpler program installed with Pathway to simplify the styles to make them compatible with Word and changes the .xhtml file to use the modified styles and add Before/After text. It then makes a copy of the xhtml file and renames

it to .doc, and then attempts to open the .doc file in Word. If you don't have Word installed on the computer but do have LibreOffice, it will open the document in LibreOffice.

4. Once the process is done, you can close the Cmd windows. Close the Word file without making any changes, and go to section 2.3. This section also explains what is wrong if Word does not show anything when it is opened.

Ideally, after step 4, you would have a Word file ready for printing. Unfortunately, that is not the case for a number of reasons, so you should not save any changes at this stage to the Word .doc file. The number of issues will depend on the dictionary data and the formatting used in FLEEx. These are some of the issues that may need to be addressed.

- Although paragraph styles are transferred from FLEEx, at this point all of the paragraphs in Word are set to the Normal style.
- Pathway produces a large number of styles, many with confusing names, because it needs to flatten the hierarchical styles in FLEEx, and it tries to maintain all of the possible settings in FLEEx dictionary configuration even though the user often uses the default formats. For example, a different character style will be used for a headword for a minor entry than used for a main entry. This is also true for styles for homograph numbers, and many other parts of the dictionary.
- Pathway is no longer supported software for further development, and development stopped before it was fully implemented and thoroughly tested, so there are a number of bugs remaining. Also, a number of changes were made in FLEEx after work ended, so these are typically not handled properly.
- The Word document at this point is single-column, and does not have any page headers with guidewords.
- All of the pictures are stored as links to external files in the .doc file. This means the word document will only include these pictures if the pictures if the "pictures" directory is available in the original location.
- Some of the styles from FLEEx do not get set correctly in Word, so you will probably need to change some of the style formatting. A strong point, though, is that the data makes heavy use of styles, so by changing the format of a style, it will change all of the occurrences in the dictionary.
- Forced line breaks, either in the FLEEx data or configuration, get dropped in the Word file.
- There may be some div sections in the Word file for special paragraphs such as senses after the first sense when senses start a new paragraph, and minor entries. Word typically has an indent for these that cannot be edited or changed in Word styles. The only way is to remove them is to edit the internal word data.
- Pathway puts out some utf-8 language tags that mess up some of the bidi commands.
- Pathway inserts hyperlinks for headwords and shows them with hard underlines. These need to be removed for a printed publication.
- Pathway mishandles FLEEx character styles that are assigned to parts of words that causes problems.
- Pathway mishandles places where there is more than one variant entry with different types.
- There are no page or division settings in FLEEx, so these need to be added in Word.

As you begin making changes to the .doc file, and even more as you start making changes in Word, keep in mind that these steps will all need to be repeated if you need to change things in FLEEx and do another export. So before making extensive changes outside of FLEEx, try to make

sure you will not need to make any further changes in FLEEx that would require a new export. The best way to do this is to set a filter in FLEEx for a small set of entries that demonstrate all of the features in the dictionary, export those entries, and see whether you can solve any remaining issues in Word. Take good notes in all of the steps you are making after exporting from Word in case you need to start over with a new export. Once you feel confident you can handle your small sample, then you can do a full export and go through the remaining steps to produce the final Word document.

The person producing the Word publication will likely need to experiment to determine how to address the problem areas. Some of the above issues are easiest to resolve by working with the original .doc file from Pathway. Other issues can be handled easier using the powerful Find and Replace function in Word that can insert text before and after styled data. Word can find any occurrence of a style, and can change that style to some other style if needed. Inside of Word. It's easy to change the page size and number of columns. It's relatively easy to set up page headers with guidewords showing the first and/or last headword on the page. However, it may not be possible to have guidewords show the current partial entry headword. It's also fairly easy to set letter headers to single column in double-column text. As in any program, placement of pictures, and dealing with location of letter headers that occur at the bottom of a page requires manual work. Some problems, such as the div markers, can only be changed in the internal .docx file. Any of the changes to the .doc file or the internal data in a .docx file requires some knowledge of XML and the way word uses XML in their internal files. Word also provides a powerful programming macro language that can be useful to those who understand that.

## 2.3 Additional processing of the .doc file before saving in Word

If the Word document opened after the Pathway conversion does not show the dictionary, it's because you are using an older version of FLEEx and/or Pathway. You can work around this problem by exiting Word without saving. Then open the .doc file in a text editor. Search for <title /> and replace it with your title (e.g., <title>This is a Title</title>). Save and close the file.

The initial state of the .doc file from Pathway is an html file with embedded styles and xhtml data using the styles. If you need to make changes, it is usually easier dealing with the file in this state. You can edit it with a text editor such as Notepad++ or ZEdit (installed in the FieldWorks 9 directory). It can also be changed with CC (SIL Consistent Changes) or other programming languages/scripts. If you save the .doc file from Word, the format is very different, and harder to work with. Here's the basic structure.

```
<html lang="utf-8" xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <title>
      <span lang="utf-8">Test - b</span>
    </title>
    <style type="text/css">body {
font-size: 10pt;
}
**many styles**
</style>
</head>
<body class="dicBody">
  <div class="letHead">
    <span class="letter" lang="fr">E e</span>
```

```

    </div>
    ** dictionary entries and letter headers ***
  </body>
</html>

```

The xhtml at this point is similar to the FLEEx xhtml export, but includes text specified in field Before/After fields in FLEEx dictionary configuration. This text is stored in the .css file, but is not supported in the html style format.

The following sections list a few changes that can be made easily at this stage, especially if you use CC or another script to make all of the changes in one pass. Assigning paragraph styles and adding forced line breaks can also be done in Word using advanced Find and Replace. But there you have to do each one separately. The utf-8 change could just be made with one search and replace in the .doc file, but can't be done in the Word UI.

If you are using a script to make the changes, it can be process from the .xhtml file to the .doc file in the export directory since initially these files have identical content.

### 2.3.1 Enable div classes to import into Paragraph styles

Paragraph styles in the .doc file are in this format

```

.entry {
  /* Rule List: 6,18,17,-1,2 */
  border-left-width: 0pt;
  border-right-width: 0pt;
  border-top-width: 0pt;
  border-bottom-width: 0pt;
  text-indent: -12pt;
  margin-left: 21pt;
  flex-line-height: 12pt;
  line-height: 12pt;
  padding-bottom: 2pt;
  padding-top: 1pt;
  clear: both;
  white-space: pre-wrap;
  font-size: 10pt;
}

```

In the xhtml data, an entry starts with <div class="entry". The goal is to have Word attach these div classes to the associated paragraph styles that are already defined. This will only work if the styles use the simple format above. If Pathway combines multiple styles together like this p.mainheadword, li.mainheadword, div.mainheadword these have to be separated into individual style rules before Word will attach them.

Word will automatically attach styles to p elements in the html file, but will not attach them for div elements without additional help. To force Word to attach a div to a paragraph style, it needs to have an extra rule that can be inserted immediately after the style name.

```

.entry {
  mso-style-name:entry;

```

In this rule, the style name must match the style into which it is inserted. This needs to be repeated for each paragraph style. Here is a CC script that processed all of the styles in one project.

```
c Enable div classes to import into paragraph styles
'.captionContent {' nl > dup ' mso-style-name:captionContent;' nl
'.entry {' nl > dup ' mso-style-name:entry;' nl
'.letHead {' nl > dup ' mso-style-name:letHead;' nl
'.minorentryvariant {' nl > dup ' mso-style-name:minorentryvariant;' nl
'.picture {' nl > dup ' mso-style-name:picture;' nl
'.sensecontent {' nl > dup ' mso-style-name:sensecontent;' nl
```

A simple way to check to make sure you've covered all of the styles is to open the .doc file in Word and search for the "Normal" paragraph style. If Normal still appears, it means something was missed in these assignments.

### 2.3.2 Insert forced breaks from style before/after text

For forced line breaks Before or After fields in FLEEx configuration (e.g., using \A in the Before field), the .doc file will have something like this for the Example Sentence field.

```
<span class="sp" xml:space="preserve"><span class="sp" xml:space="preserve">
</span></span><span class="example" lang="trw">
```

There is a CRLF in the sp span, but Word does not interpret this as a forced line break (Shift+Return in Word UI). Word needs to have <br/> at this point. The above xml, including the CRLF could be changed by search and replace or in a script to replace it with <br/>. Alternatively, you could search for the styles that need to be changed and add the break as in this CC rule.

```
'<span class="example" > ' <br/>' dup
```

### 2.3.3 Replace utf-8 "language"

Pathway inserts lang="utf-8" in various places including in the <html begin tag. UTF-8 is a type of encoding for any language, but is not really a language. In a bidi dictionary in Word, some strange formatting occurred in some places, and it turned out that the use of utf-8 as a language in the html begin statement was the cause of the problem. After changing all lang="utf-8" to lang="en" the problems went away. It would be wise to make this change throughout the .doc file. Here's a CC rule to do this.

```
'lang="utf-8" > 'lang="en"
```

## 2.4 Preparing a Word dictionary for publication

After opening the .doc file in word and verifying that all paragraph styles are assigned and any other changes made to the .doc file have been made, then save the file to a normal Word .docx file. At this point advanced Find and Replace can be used to make additional changes, or manual edits can be made.

### 2.4.1 Page layout

The dictionary that comes from FLEEx does not have any page layout information. It just contains the dictionary entries and letter headers. The first thing that needs to be done in Word is to set the page size, orientation, margins, and any other page or book type settings from the Layout tab. If the dictionary is to be double-column, use Ctrl+A to highlight the entire document, then in Layout...Columns choose "Two".

## 2.4.2 Removing hyperlinks

At this point the headwords and links to other entries contain hyperlinks and they have hard underlines to indicate they are hyperlinks. These have no purpose in a printed dictionary, so they should be removed. To do this, use Ctrl+A to highlight the entire document, then use Ctrl+Shift+F9 to remove all hyperlinks.

This still leaves hyperlink underlines which are hard coded, rather than being part of a style. To remove these, use Ctrl+A to highlight the entire document, then in the Home tab, click the U underline button to force everything to underline. Then click it again to remove all of the underlines. If there are any underlines in FLEEx entries, these should be in a style which should not be affected by this procedure.

## 2.4.3 Embedding pictures in the .docx file.

When FLEEx exports to XHTML, any pictures are copied to a pictures directory in the xhtml folder, and external links are used in entry data to the files in this folder using relative links.

```
<div class="picture">
```

When Pathway creates a Word file, and the file is opened in Word, the pictures are converted to links to external files using the original location of the file on the hard drive. If the Word document is passed to other users, the pictures will not be available. Section 3.43.4 describes how to change these external links to embedded pictures in the .docx file. Then when the file is passed to someone else, they still see the pictures in the file.

## 2.4.4 Adjusting styles

Paragraph and Character styles are transferred from FLEEx to Word to some degree, but Pathway generates different names from the ones used in FLEEx. The dictionary data uses these styles extensively. See section 3.1 for basic information on working with styles in Word. See section 3.33.3.13.3.1 for ways to search for styles in word, section 3.3.2 for replacing styles, and section 3.3.53.3.5 for adding text before or after styles.

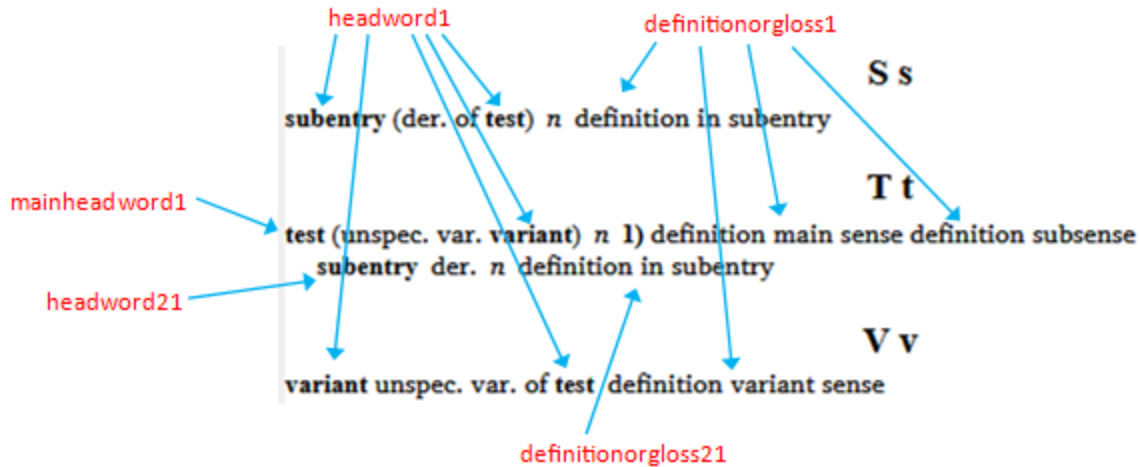
One of the complications with styles coming from FLEEx via Pathway is the large number of styles that end up in the word document.

Pathway provides many paragraph styles for FLEEx configurations where this may be important. For example, entry, headword, headword2, mainheadword, minorentrycomplex, minorentryvariant are all paragraph styles that could be configured differently, so in some contexts you may want to keep these distinctions, but typically you might want to merge some of these to reduce the complexity.

Pathway provides many character styles for FLEEx configuration where this may be important. For example, definitionorgloss1, definitionorgloss21, headword1, headword21, mainheadword1 are all character styles that could be configured differently, so in some contexts you may want to keep the distinction, but typically you might just merge all of these to definitionorgloss1.

Here's a diagram that shows how this works for a case where default styles were used in FLEEx.





From this you can see that

- mainheadword1 is used for headwords of main entries.
- headword21 is used for headwords of subentries in a main entry
- headword1 is used for all other places headwords are used
- definitionorgloss21 is used for definitions in subentries in a main entry
- definitionorgloss1 is used for all other definitions

It's unfortunate that these style names do not match the style names in FLEEx. Word Find and Replace provides an easy way to change one style to another. It would be easy to change definitionorgloss21 to definitionorgloss1 if they are all the same format. If you don't merge the style names, then when you make a change to one style, you need to find all of the other styles and make the same change to those.

You will probably need to make some changes to styles to get the desired output. Watch for places where there is more than one style that needs to be changed. This includes styles for homograph numbers, sense numbers, etc.

One other thing to be aware of if you need to export again from FieldWorks, is that the style names could change in new exports if you've made changes to style names in FLEEx that would cause new collisions with existing names. So, if you have a process that depends on certain style names and you do another export, there is a slight chance of having different style names.

If you would like to get rid of unused styles, see section 3.63.6. **Error! Reference source not found.**

## 2.4.5 Add page headers with guidewords

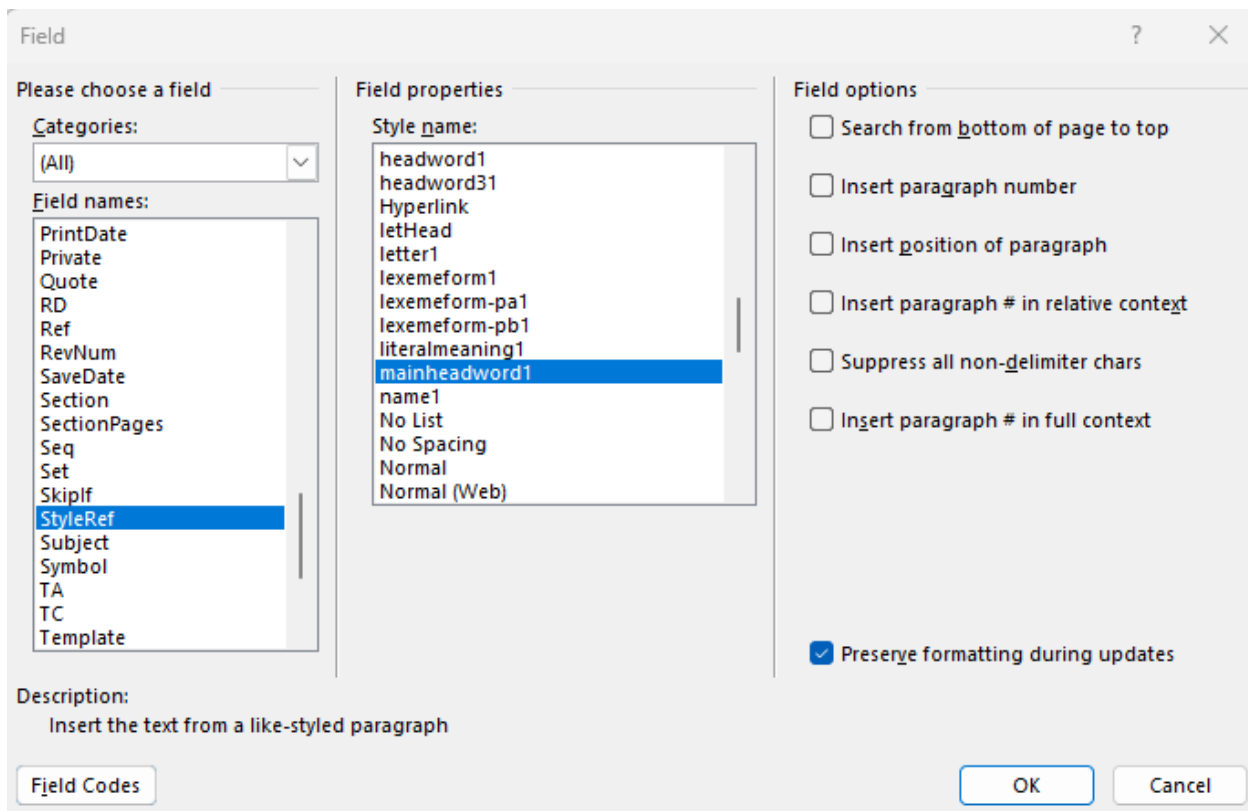
It's fairly easy to set up page headers and footers in Word to handle various needs. To edit headers, choose Insert...Header...Edit Header. This opens a header section with a Design tab where you can add guidewords, page numbers, text, etc., and format the sections as desired. The first tab is a centered tab and the next tab is a right tab. When done editing the header, use Design...Close Header and Footer.

In the header tab you can select whether to use a different header for the first page, and whether you want different even and odd page headers.

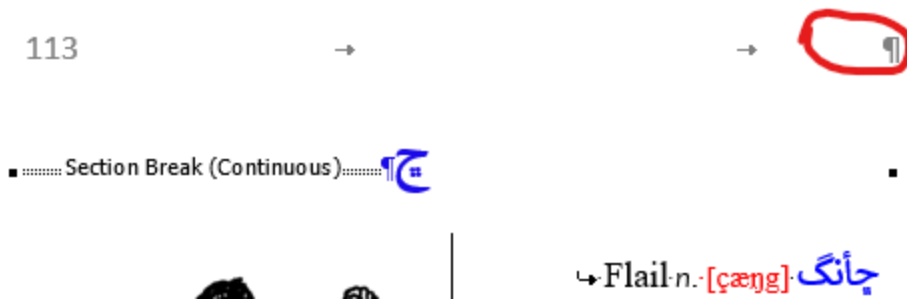
To insert a page number, use Design...Page Number...Current Position...Plain Number, or whatever version you desire.

Word provides an easy way to insert text from the first style on a page and/or the last style on a page. This provides a way to show dictionary headwords for the current page. There are some considerations in doing this. First, as noted in section 2.4.4, the default style for headwords provided by Pathway are mainheadword1 for main entries, but headword1 for minor entry headwords. When setting up guidewords you can only use one style, so assuming you want to allow minor entries in the guidewords, you'll need to change those styles to mainheadword1. Unfortunately, if you change all of them, you'll also change variant entries and main entries for minor entries to the same style, which will likely give undesirable results. So, you'll need to be more selective than a global style change to get desirable results. The other issue is what to do when a page begins with a continuation of a longer entry from the previous page. It appears that Word cannot handle this in the normal page header settings. The only easy option is to include headwords from the current page.

To print a headword in the header, use Design...Quick Parts...Field. In the Field names, choose StyleRef, and in the Field properties, choose the headword character style (e.g., mainheadword1). If you want the last headword on the page, check the "Search from bottom of page to top" option.



You might discover that after setting letter headers, if the header comes at the top of a page, the first headword is not displayed. Here is an example



The most likely problem with this is that the Section Break (Continuous) is set to the mainheadword1 style, so it has no text, and it's the first occurrence of the style on the page, so it doesn't show anything. To solve this, click Section Break (Continuous) and set the style to letter1. The guideword will then appear. If you do not have File...Options...Display..."Show all formatting marks" checked, you won't see this text, but if you click to the left of the letter header, you'll probably see that the style is set to mainheadword1.

#### 2.4.6 Convert letter headers to single column

When using two columns for your dictionary text, it's usually desirable to have the letter headers span both columns. Word does not provide a style for doing this as InDesign provides, but there is a solution that works reasonably well that just requires making a change for each letter header. Pathway normally uses the letter1 style for letter headers. By searching for this style in the Advanced Find dialog, you can quickly find the next letter header. You can then choose Layout...Columns...One, which will make a single column section break around the letter header so that it will span both columns. If a letter header occurs at the bottom of a page, you can insert Shift+Returns to push the letter header to the top of the next page. Word does pretty well at repositioning letter headers when you make changes earlier in the document. If things get too bad, it may be easiest to remove the letter header section breaks and then go through the process of resetting the letter headers again. Someone who does this a lot may benefit in setting up a Word macro to do it in one step.

#### 2.4.7 Editing the internal Word docx file.

See section 3.8 for basic information on modifying the internal Word document.

Where this process might be necessary is a situation where Pathway inserted some divisions following additional senses when senses each start a new paragraph, and where it added a division after minor entries. The problem was that it produced an indent that could not be removed in the styles or hard formatting. Once the division was removed in the internal file, then it was possible to set the indents as desired. The internal XML had this data

```
<w:pStyle w:val="minoreentryvariant"/><w:divId w:val="1917278976"/>
<w:pStyle w:val="sensecontent"/><w:divId w:val="254558700"/>
```

The yellow portions are the divisions that caused this problem. The number in each instance was different, so a simple search and replace would not work. However, with Notepad++ with regular expression enabled, you can search for (<w:pStyle w:val="sensecontent"/>)(<w:divId w:val="\d+?"/>) and replace it with \$1. Replace All removes all of them following sensecontent. This solved the problem with senses after the first sense. Then repeating the process for minoreentryvariant took care of the minor entry problems.

## 2.4.8 Dictionary reversal indices

Pathway optionally includes one reversal index with the main dictionary when doing a normal Pathway export. When you have more than one index, and when you are using Pathway to create a Word document, you need to export each index separately, process the index as described above, and then add it to the master dictionary document.

It's usually good to clean up indices before publication. To do this, go to Lexicon...Bulk Edit Reversal Entries. It's possible that in various editing operations on reversals, you will have unused reversal entries, either with missing reversal forms or missing referenced senses. These should be deleted using the Delete tab. Check for missing reversal forms by filtering the Reversal Form column for Blanks. If there are any blank entries, select Reversal Entries (Rows) as the Item to delete, make sure they are all checked, then click the Delete button. Clear the Reversal Form filter and filter the Referenced Senses column for Blanks. Again, if any show up, make sure they are all checked, then click the Delete button. When done, clear the filter. This process needs to be done for each reversal language.

In Bulk Edit Reversal Entries you can set a filter to limit the reversal entries to export, and you can also make sure the reversal index is sorted correctly. If there are problems with the sorting or letter headers in the reversal, you need to adjust the Sorting tab of the reversal writing system. Normally you'll want to sort on Reversal Form and not have any filter active. Whatever is set in Bulk Edit Reversal Entries will determine what shows in the reversal index for that language.

When this is set correctly, go to Lexicon...Reversal Indexes In the blue title bar, select the index language in the chooser on the left side, and select the desired view in the chooser on the right side. Then choose File...Export...Reversal Index Web page (XHTML) and export it into an empty folder on your machine. Repeat this for each index you want to include with the Word dictionary.

After exporting the indices exported, use the process starting at step 3 in section 2.2 to produce Word files for each index. Then they can be combined into a master Word file for the final publication.

## 2.4.9 Dictionary front matter and back matter

Use Word's normal editing capabilities to add any front matter (title pages, table of contents, grammar sketches, description of dictionary, etc.) or back matter (appendices, etc.)

# 3 Useful features in Word

## 3.1 Basics of Styles in Word

There are two common types of styles in Word: Character, and Paragraph. Paragraph styles let you control paragraph indents, spacing between lines, alignment, bullets, numbering, and line and page break properties of paragraphs. Character styles let you choose the font, font face, font size, color, background color, subscript, superscript, underlines, etc. The advantage of using styles is consistency and ease of use. By modifying a style, you will affect all of the places that style is used. It's generally much better making changes to styles than doing hard formatting in your document.

When working on the dictionary in Word, it's very helpful to have the stylesheet open beside the document. If you don't have the stylesheet open, you can open it with Alt+Ctrl+Shift+S, or by clicking the button in the lower right of the Styles pane in the Home tab. Once open, when you click in a word in the dictionary, the associated style is selected in the style sheet.

If you are interested in Paragraph styles instead of character styles, it works best to have paragraph marks displayed in your document. This is controlled in File...Options...Display. With paragraph marks showing, you can select just the paragraph mark, and then the stylesheet will select that paragraph style.

A style is applied by selecting the target text, then clicking the style in the style sheet. Paragraph styles can be applied by clicking the style from anywhere in the paragraph. Paragraph styles also contain character formatting that will be applied to any text in the paragraph that is not controlled by some other character style.

To modify a style, from the stylesheet, right-click the style and choose Modify. This brings up the Modify Style dialog where you can change the name and basic characteristics of the style. For more complete control, you can click the Format menu in the lower left, then choose Font, Paragraph, Tabs, etc., to view and modify more specific features in the style. When you click OK to the Modify Style dialog, all text using that style will update to the new style.

## 3.2 Preventing styles from changing: dynamic updating

Word provides dynamic update of many styles. When enabled, Word will change the definition of a style when you apply explicit formatting to something in the document. Especially when multiple people are working on the same document, this can cause styles to unexpectedly change, which is typically undesirable. To avoid this, make sure "Automatically update" is unchecked in the Modify Style dialog. To disable this in all styles, you can use this macro:<sup>1</sup> See section 3.7 for information on using macros.

```
Sub RemoveAutoUpdate()  
  Dim s As Style  
  For Each s In ActiveDocument.Styles  
    If s.Type = wdStyleTypeParagraph Then  
      s.AutomaticallyUpdate = False  
    End If  
  Next s  
End Sub
```

## 3.3 Advanced Find and Replace

Word provides powerful find and replace capabilities. It doesn't provide regular expression searching, but it does provide wildcards for some options. There are three ways to get to the Advanced Find and Replace dialog,

- Home...Find...Advanced Find.
- In the Navigation section, click the chooser at the right of the Search edit box, and choose Advanced Find
- Ctrl+H opens the Advanced Find and Replace dialog in the Replace tab. Click the Find tab if not interested in replacing.

Once in the Advanced Find and Replace dialog, there are 3 tabs: Find, Replace, and Go To. The default view shows minimal capabilities, but a More button opens up many more features. The rest of this discussion assumes the full features are available.

### 3.3.1 Finding styles

One very useful feature is finding styles in a document. To do this, leave the Find edit box empty, but at the bottom of the dialog choose Format...Style, and then select the target style and click OK. This adds a Format label under the Find box that tells which style is selected. When you click Find Next, it will find and highlight the next text in the target style. This can be very useful in finding the data represented by styles that Pathway created. You can remove the style from the Find or Replace box by choosing Format...Style...(no style), or you can click the “No Formatting” button at the bottom of the dialog.

### 3.3.2 Replacing styles

You may find it useful to merge styles provided by Pathway to make them more consistent. For example, Pathway gives different styles for definitions depending on whether they are in a main entry (definitionorgloss1) or a subentry within an entry (definitionorgloss21). Typically, these will be the same, so it’s nice to change one style rather than two or more.

To change a style using the Find tab of the Find and Replace dialog (Ctrl+H), make sure the Find and Replace edit boxes are empty, click the Find edit box, and use Format...Style to select the source style, then click the Replace edit box and use Format...Style to select the destination style. Then click Replace All to change every occurrence of the source style to the destination style.

To turn off styles in the Find and Replace boxes, you need to use Format...Style and select (no style) at the top of the list, or click the “No Formatting” button at the bottom of the Find and Replace dialog.

### 3.3.3 Special codes

Word provides a set of special codes starting with a caret (circumflex) that can be very useful. If you select an item from the Special menu at the bottom of the Find and Replace dialog, it will add the specified code at the cursor in the Find or Replace edit box. You can also type the codes directly. Here is a list of the special codes:

<b>^p</b>	Paragraph Character	
<b>^t</b>	Tab Character	
<b>^?</b>	Any Character	
<b>^#</b>	Any Digit	
<b>^\$</b>	Any Letter	
<b>^^</b>	Caret Character	Find caret or circumflex (^) code point
<b>^n</b>	Column Break	
<b>^+</b>	Em Dash	
<b>^=</b>	En Dash	
<b>^e</b>	Endnote Mark	
<b>^d</b>	Field	‘View Fields Codes’ must be on. See below.
<b>^f</b>	Footnote Mark	

<b>^g</b>	Graphic	Any picture or graphic.
<b>^l</b>	Manual Line Break	Equivalent to typing Shift + Enter
<b>^m</b>	Manual Page Break	Equivalent to typing Ctrl + Enter
<b>^~</b>	Nonbreaking Hyphen	
<b>^s</b>	Nonbreaking Space	
<b>^-</b>	Optional Hyphen	
<b>^b</b>	Section Break	
<b>^w</b>	White Space	Space, tabs, etc.
<b>^%</b>	Section Character	The § symbol. Only used in Replace box.
<b>^v</b>	Paragraph Character	The ¶ symbol. Only used in Replace box.
<b>^r</b>	Right-to-left Mark	
<b>^h</b>	Left-to-right Mark	
<b>^y</b>	Zero Width Joiner	
<b>^o</b>	Zero Width Non-Joiner	
<b>^uN</b>	Unicode character	N is a decimal number (only in Find box) e.g. to find Unicode line separator U+2028, ^u8232
<b>^&amp;</b>	Insert Find text	This only works in the Replace field and inserts the target of the find operation.

To show field codes on a page, use Alt+F9. To show field codes for a single field, click in the field and use Shift+F9. To show field codes in all documents, use File...Options...Advanced...Show document content...Show field codes instead of their values. When you want to stop showing field codes in all documents, you need to uncheck the Show field codes option.

### 3.3.4 Wild cards

Although Word does not provide regular expression search and replace, they do provide some of this capability when the “Use wildcards” option is checked. When checked, the Special button at the bottom of the Find and Replace dialog offers a limited set of options.

In the Find box:

<b>?</b>	Any Character	
<b>[ - ]</b>	Character in a Range	[abc] a, b, or c; [a-z] letters a through z
<b>&lt;</b>	Beginning of word	
<b>&gt;</b>	End of word	
<b>( )</b>	Expression	Allows nesting and used in replacement
<b>[ ! ]</b>	Not	[!aeiou] not a vowel
<b>{ , }</b>	Num occurrences	^p{3} finds 3 paragraph marks; a following comma means at least that number. ^p{3,} finds 3 or more paragraph marks; {n,n} min and max ^p{2,3} finds 2 or 3 paragraph marks
<b>@</b>	Previous one or more	^p@^t finds one or more paragraph marks before tab
<b>*</b>	Zero or more	
<b>^t</b>	Tab character	
<b>^^</b>	Caret character	

<b>^n</b>	Column Break	
<b>^+</b>	Em Dash	
<b>^=</b>	En Dash	
<b>^g</b>	Graphic	
<b>^l</b>	Manual Line Break	
<b>^m</b>	Manual Page / Section Break	
<b>^~</b>	Nonbreaking Hyphen	
<b>^s</b>	Nonbreaking Space	
<b>^-</b>	Optional Hyphen	
<b>\wildcard</b>	Wildcard character in text	\* for *, \> for >, etc.
<b>^N</b>	ASCII code	^013 or ^13 is ASCII decimal code 13 which is the same as ^p when not in wildcard mode. It doesn't extend to Unicode values (only 1-127)

In the Replace box:

<b>\n</b>	Find what expression	Group number in find expressions. See below
<b>^p</b>	Paragraph mark	
<b>^t</b>	Tab character	
<b>^^</b>	Caret character	
<b>^%</b>	Section Character	
<b>^v</b>	Paragraph Character	
<b>^c</b>	Clipboard contents	
<b>^n</b>	Column Break	
<b>^+</b>	Em Dash	
<b>^=</b>	En Dash	
<b>^&amp;</b>	Find what text	Insert the target text from the find operation
<b>^l</b>	Manual Line Break	
<b>^m</b>	Manual Page Break	
<b>^~</b>	Nonbreaking Hyphen	
<b>^s</b>	Nonbreaking Space	
<b>^-</b>	Optional Hyphen	
<b>^N</b>	ASCII code	^013 or ^13 is ASCII decimal code 13 which is the same as ^p when not in wildcard mode. It doesn't extend to Unicode values (only 1-127)

Here is an example demonstrating the “Find what expression”. Suppose we have a list of names

John Doe  
Sally Smith

And we want to change this to

Doe, John  
Smith, Sally

In the Find box, type

`([A-z]@) ([A-z]@)^13`

This starts with an expression in parentheses with a sequence of one or more upper or lower case



letters ([A-z]@) followed by a space then a second expression of one or more upper or lowercase letters ([A-z]@). When in the “Use wildcards” mode, we can’t use ^p for a paragraph mark, so we use the ASCII code syntax instead, which is ^13 since the code for Carriage Return (CR) is decimal 13.

In the Replace box, type

\2, \1^p

The replacement is the content from the second expression in the find string \2, then a comma and a space, then the first expression in the find string \1, then a CR ^p. Note that ^p is valid in the Replace box and also maintains the style/formatting information of the original paragraph, while using ^13 would have lost the paragraph formatting. The result after clicking “Replace All” are the original names in reverse order.

### 3.3.5 Adding text before or after a style

Using a combination of the Advanced Find features, we can easily add text before or after a style throughout a document. Suppose the import missed putting a Return before an example sentence. To do this, set up the Find and Replace boxes with the example1 character style. Leave the Find box empty, and put ^|^& in the replace box. Then Replace All.

The screenshot shows the 'Find and Replace' dialog box with the following settings:

- Find what:** (empty text box)
- Format:** Style: example1
- Replace with:** ^|^&
- Format:** Style: example1

With an empty Find box, and example1 as the style, the Find operation will find and highlight the next example sentence text (everything in the example1 style). The Replace operation will keep the example1 style for the output. The content will be a forced line break (Shift+Enter) from ^|, and then the content of the find operation ^& which would be the original example text.

One side-effect of this operation is that the line break character is also part of the example1 style. If that becomes a problem, it’s challenging getting rid of that. If you highlight the Shift+Return character and use Ctrl+Space, Word removes the style as desired. When you check the style after that, it will be the entry paragraph style. You could search for the ^| with the example1 style, and replace it with ^| in some other style. But unfortunately, using (no style) should do the trick, but Word doesn’t change the style in this case. Also, using the entry style for the replacement should work, but again, Word doesn’t change the style. It will only change the style if you change it to some other character style. One option would be to create an entry1 character style and assign it to that. The only other option may be to remove the character style in the internal XML as described in section 2.4.7.

## 3.4 Saving pictures in the Word document.

Word can store pictures as links to external files, or embedded in the .docx file.

With external links, each picture is given a new guid as an id. When this gets saved to .docx, internally, word\document.xml has <a:blip r:link="rld5">, and word\\_rels\document.xml.rels has

```
<Relationship Id="rld5"
Type="http://schemas.openxmlformats.org/officeDocument/2006/relationships/image"
Target="file:///W:\Users\username\Desktop\yy\pictures\x.png" TargetMode="External"/>.
```

Note the path is a fixed path on the machine that created the .docx file. So, when the Word document is passed to other users, the pictures will not be available.

The solution for this is to embed the pictures in the .docx file, which is what usually happens when you add a picture to a Word document. To embed the pictures, you can run a Word macro<sup>ii</sup>. See section 3.7 for information on using macros. You can then save your Word document to save the changes.<sup>iii</sup>

```
Sub convert_all_inline_shapes()
'Updated by Extendoffice 20180906
Dim xlShape As InlineShape
For Each xlShape In ActiveDocument.InlineShapes
With xlShape
If .Type = wdInlineShapeLinkedPicture Then
.LinkFormat.SavePictureWithDocument = True
.LinkFormat.BreakLink
End If
End With
Next
End Sub
```

The resulting .docx file will now include the picture so that it can be copied to anyone and they can open the .docx file and see the pictures. Internally, word\document.xml has

```
<a:blip r:embed="rld4">, word\_rels\document.xml.rels contains <Relationship Id="rld4"
Type="http://schemas.openxmlformats.org/officeDocument/2006/relationships/image"
Target="media/image1.jpg"/>
```

and word\media has image1.jpg. The file has been converted to jpg internally and has a new name. But the picture is now stored in the .docx file. The document.xml file still has the original picture name with a full path to the original location, but it is only in a descr element, and not used for display purposes.

### 3.5 Removing hyperlinks in Word

If you need to remove hyperlinks in a Word document, use Ctrl+A to highlight the entire document, then use Ctrl+Shift+F9 to remove all hyperlinks.

### 3.6 Removing unused styles in Word

If you would like to get rid of unused styles, there is a macro available that will do this. However, before doing this, you should be careful that you no longer need the unused styles. Of course, you can always add new styles if needed. See section 3.7 for information on using macros. Here is the script for doing this.<sup>iv</sup>

```
Sub DeleteUnusedStyles()
Dim oStyle As Style
For Each oStyle In ActiveDocument.Styles
'Only check out non-built-in styles
```

```
If oStyle.BuiltIn = False Then
  With ActiveDocument.Content.Find
    .ClearFormatting
    .Style = oStyle.NameLocal
    .Execute FindText:="", Format:=True
    If .Found = False Then oStyle.Delete
  End With
End If
Next oStyle
End Sub
```

### 3.7 Running macros in Word

Word provides a very powerful Visual Basic programming language to write macros, or scripts, that can be run automatically through the document. You can record sequences of commands as a macro, and then run that later to repeat those steps. Useful macros can often be found on the Internet.

If you have a macro script, the first thing you need to do is store that script in Word. One way to do this is to press Alt+F11 to open the VB window. Paste the macro VB code in the window. If you want to run it immediately, with the cursor at the beginning of the macro code, press F5 to run the script. Use File...“Close and return to Microsoft Word” to close the VB window and go back to your document.

Once a macro script is stored in Word, it can be run in the future by using View...Macros...View Macros, click the macro name and click Run to run it.

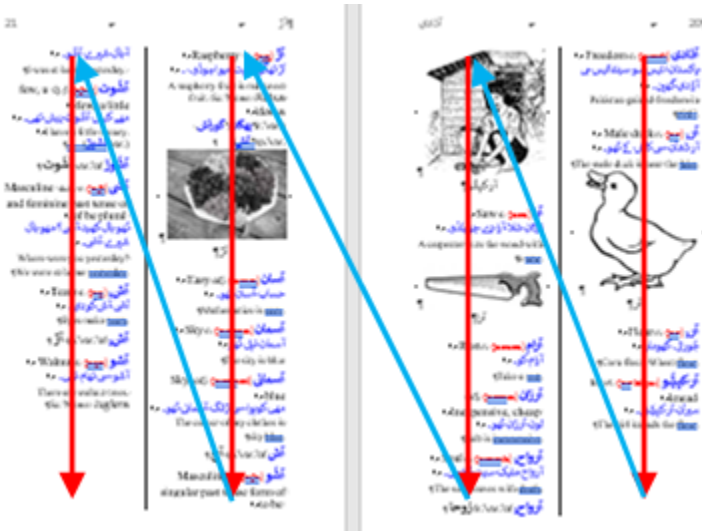
### 3.8 Editing the internal Word docx file.

A Word .docx file is actually a renamed zip file that has several folders and files internally. With the free 7-Zip program (<https://7-zip.org/>) you can right-click a .docx file and choose 7-Zip...Open archive. You can also right-click the file and choose 7-Zip...Extract to ... which will unzip the file to a directory with the name of the .docx file. In this directory you can explore the various files. The file that contains the text in the document is word\document.xml. This file contains word XML data. If you know what you are doing, you can modify this file. To get it back in the .docx file, zipping the directory will probably not work. But you can open the .docx file again in 7-Zip and then drag the modified document.xml file into the word directory. 7-Zip will save the file. You can then close 7-Zip and open the file in Word. Of course, you should always do this on a copy of the original file in case you mess up the file in a way it can no longer be opened.

## 4 Bidirectional dictionaries in Word

Bidirectional (bidi) dictionaries have data in both left-to-right (LTR) and right-to-left (RTL) scripts. A standard bidirectional algorithm (see 4.3) is used to lay out a paragraph with mixed directions. Entries can be laid out in a LTR or RTL page layout. The direction chosen is typically based on the primary reading orientation of the intended audience. If the language of headwords is RTL and the page layout is LTR, its harder to locate headwords since the part of the headword at the margin is the end of the RTL word. A RTL page layout is easier to read in this respect.

In a RTL publication, text is laid out as shown in this sample:



Text is read from right-to-left within a line, and reading continues down the right side of the page to the next column, and continues across a page spread from right to left. Page numbers increase from right to left. RTL (e.g., Arabic) books are opened from the opposite side as they are for a LTR (e.g., English) book.

#### 4.1 Enabling right-to-left publications in Word

Word does a reasonably good job of handling right-to-left publications with page numbers, headers, and columns all flowing RTL instead of LTR. However, it's not at all obvious how you can get Word to work this way. It turns out, that Word provides some extra options if you have a right-to-left language installed in Office.

These instructions are for Office 2019 on Windows 11, so may vary somewhat with other versions. To install the language, before opening Word, use the Windows start icon to find Microsoft Office Tools, and under that, open Office Language Preferences. In this program, you need to have at least one RTL language installed as an Editing Language. If you do not have one, and you do not have a preference for which RTL language to install, in the "Add additional editing languages" chooser, select Arabic (Saudi Arabia) and then click Add. At this point it won't be enabled, so click the "Not enabled" link. This will take you to the Windows Language & region dialog where you will need to click "Add a language". After selecting the language in the list, click Next, and select the Language pack option. Be careful not to check "Set as my Windows display language". Click Install to install the language pack. You do not need to install the other options such as speech recognition and proofing options. You can then click OK to close the Office Language Preferences dialog. Now Word will have options to produce RTL books.

For RTL book layout, go to File...Options...Advanced, there are some new features under Editing Options, and Show Document Content. Under Show Document Content, Set Document View to Right-to-left. At this point Word will properly lay out books in RTL mode. Unfortunately, this is a global Word setting, so if you have a mix of LTR and RTL documents, you'll have to change this each time you switch to another direction.

## 4.2 Basic RTL layout in FLEx and Word

When the primary language in a publication is right-to-left, here are some hints on how to set up the language and configuration to handle this.

### 4.2.1 Settings in FLEx

These are the basic settings you need in FLEx for an RTL publication

- In Format...Set up Vernacular Writing System, make sure “Right-to-left” is checked in the General tab.
- In Format...Styles...Normal...Paragraph, set Direction to Right-To-Left, and set Alignment to Right.

### 4.2.2 Settings in Word

These are the basic settings you need in Word for an RTL publication.

- In File...Options...Advanced...Show Document Content, set Document View to Right-to-left. If this option isn't available, see section 4.1.
- With the style sheet showing (Alt+Ctrl+Shift+S), right-click the Normal style and choose Modify, then at the bottom left of the Modify Style dialog, choose Format...Paragraph. In the Paragraph dialog set Alignment: to Right and check Direction: Right-to-left.
- In the Layout tab, click the icon at the lower right of the Paragraph section, and in the Paragraph dialog set Alignment: to Right and check Direction: Right-to-left.
- In the Layout tab, click the icon at the lower right of the Page Setup section, and in the dialog that comes up, set Section direction: to Right-to-left.

## 4.3 Bidirectional algorithm

Unicode provides guidelines on how bidirectional (mixture of left-to-right and right-to-left text) should be displayed in a paragraph<sup>1</sup>. For a simple example, take this simple dictionary entry with a headword, a definition, an example sentence, and a translation of the example.

آشی

Masculine and feminine past tense of “be” plural.

تھو بال کھید آشی؟ مہو بال شیرے آشی۔

Where were you yesterday? We were at home yesterday.

It's pretty easy to read when each field is on a line by itself. But when it is put together in a paragraph, things get a little more complicated.

آشی Masculine and feminine past tense of “be” plural. آشی  
 مہو بال شیرے آشی۔ Where were you yesterday? We were at home yesterday.

<sup>1</sup> Unicode Bidirectional Algorithm <https://unicode.org/reports/tr9/>

The algorithm lays out text RTL and puts what it can in each line. In the top line after the description, the example only partly fits on the line, so the rest of the example goes on the next line, starting at the right. The translation fills out the rest of the second line, but the final part overflows to the third line. Now if this gets printed in double-column, it gets more interesting.

Masculine and feminine past tense of "be" اسی  
 Where اسی۔ مہو بال شیرے اسی۔ plural.  
 were you yesterday? We were at home  
 yesterday.

In this case the definition flows on to the second line. The example fits on the second line. But the translation starts on the second line, fills up the third line, and finishes on the fourth line. If this were printed in three columns, the result is this.

Masculine and feminine اسی  
 اسی past tense of "be" plural.  
 بال شیرے اسی۔ مہو بال شیرے اسی۔  
 Where were you yesterday?  
 We were at home yesterday.

The definition fits on the first and second lines. The example starts on the second line and ends on the third line. The translation goes on the third and fourth lines.

There are times the default algorithm needs to be overridden to produce desirable results. Unicode provides several code points to assist in handling special cases. These code points are summarized in this table.

Abbr.	Code	Name	Description
LRM	200E	Left-to-Right Mark	Left-to-right zero-width character
RLM	200F	Right-to-Left Mark	Right-to-left zero-width non-Arabic character
ALM	061C	Arabic Letter Mark	Right-to-left zero-width Arabic character
LRE	202A	Left-to-Right Embedding	Treat the following text as embedded left-to-right.
RLE	202B	Right-to-Left Embedding	Treat the following text as embedded right-to-left.
LRO	202D	Left-to-Right Override	Force following characters to be treated as strong left-to-right characters.

RLO	202E	Right-to-Left Override	Force following characters to be treated as strong right-to-left characters.
PDF	202C	Pop Directional Formatting	End the scope of the last LRE, RLE, RLO, or LRO.
LRI	2066	Left-to-Right Isolate	Treat the following text as isolated and left-to-right. (LRI will not work in Word.)
RLI	2067	Right-to-Left Isolate	Treat the following text as isolated and right-to-left. (RLI will not work in Word.)
FSI	2068	First Strong Isolate	Treat the following text as isolated and in the direction of its first strong directional character that is not inside a nested isolate. (FSI will not work in Word.)
PDI	2069	Pop Directional Isolate	End the scope of the last LRI, RLI, or FSI. (PDI will not work in Word.)
NBSP	00A0	No-Break Space	A space that will not allow a line break. (This is not a bidi code point, but is useful in various places.)

#### 4.4 Formatting bidirectional entries in FLEx

Bidirectional dictionaries provide numerous challenges in getting a reasonable result. Some of these issues will be explored in this section. FLEx supports all of the special codes summarized in the table in section 4.3. These codes could be inserted in the middle of fields, but most fields are a single writing system so they are not needed. However, when Before and After text is added in dictionary configuration, this text does not have a writing system. When mixed with bidirectional fields, getting the right results can be challenging. It's further challenging because programs do not always treat the bidi code points the same. One obvious case is that Word does not support LRI, RLI, FSI, and PDI, so even though FLEx can use these, if you plan to export to Word, they cannot be used. So even though you can get a perfect display in FLEx, when it gets exported to Word, you may find that Word doesn't follow the same rules. It may take a lot of trial and error to get the desired formatting in FLEx that also looks the same in Word. What works best is to pick a small set of entries that demonstrate all of the field and paragraph possibilities and then export these to Word. In places that Word doesn't work, by trying some different bidi commands, it's usually possible to come up with a set that will work in both places.

In dictionary configuration in FLEx, the code points in the bidi chart in section 4.3 can be entered in Before, Between, and After configuration for each field. To enter the codes, enclose the abbreviation from the chart in square brackets. For example, the RLE code would be entered in the configuration as [RLE]. In the underlying Word file, this would be Unicode character U+202B.

Suppose we have a simple entry that includes a headword, a pronunciation, a part of speech, and a definition. The goal is to have it print like this in Word

Junction, intersection n. [du:jəŋga] دُوجنګا

To put a space after the category and square brackets around the pronunciation, we would normally use this, where the middle dot is really a space in configuration.

Gram Info (Abbr): After: .·

Lexeme Form (IPA): Before: [ After: ]·

Junction, intersection .n ]du:jəŋga[ دُوجنګا

Two problems are obvious immediately in FLEx. The period is on the wrong side of the category, and the square brackets are backwards. The basic problem is that Before and After text does not have a writing system. So, when FLEx lays out the pieces going right-to-left, we have the headword, then the open square bracket, the pronunciation, the closing square bracket, the category, the period, and then the definition. We could get the desired output by putting the period in Before text and putting in backward square brackets in the Before and After. But this is technically not correct, and will cause other problems along the way.

One option is to use left-to-right isolates around the grammar info and the lexeme form.

Gram Info (Abbr): Before: [LRI] After: .[PDI]·

Lexeme Form (IPA): Before: [LRI][ After: ][PDI]·

Junction, intersection n. [du:jəŋga] دُوجنګا

This works perfectly in FLEx. Unfortunately, Word does not support the isolates. It actually gets the period and square brackets right, but it displays the isolate codes with undesirable glyphs. It looks like this:

Junction, intersection [LRI][du:jəŋga][PDI] [LRI]n [PDI] دُوجنګا

We can test the embedding codes to see how that works.

Gram Info (Abbr): Before: [LRE] After: .[PDF]·

Lexeme Form (IPA): Before: [LRE][ After: ][PDF]·

Junction, intersection [du:jəŋga] n. دُوجنګا

This gets the square brackets and period working correctly in FLEx, but now there is a positioning problem. The pronunciation is an entry field and should follow the headword. But in this case the category is following the headword. The bidi algorithm sees the headword and category as one LTR string, which ends up with the wrong order.

The way to solve this is to use the Arabic Letter Mark which prevents the bidi algorithm from putting the pronunciation and category together.

Gram Info (Abbr): Before: [LRE] After: .[PDF]·

Lexeme Form (IPA): Before: [LRE][ After: ][PDF][ALM]·

Junction, intersection n. [du:jəŋga] دُوجنګا

Now FLEx is giving us what we want. But, there is new problem when we export to Word.

n. Junction, intersection [du:jəŋga] دُوجنګا



The bidi algorithm in Word works differently than FLEx, so we bump into a similar problem where Word is seeing the category and definition as one LTR string, so puts the category in the wrong position.

We can solve this by adding an Arabic Letter Mark to category as we did to the pronunciation. There is also an extra space generated with a trailing space in the category After text, so that has been removed.

Gram Info (Abbr): Before: [LRE] After: .[PDF][ALM]

Lexeme Form (IPA): Before: [LRE][ After: ] [PDF][ALM].

Junction, intersection n. [du:jəŋga] دُوجنْگا

FLEx still shows what we want. Now Word also shows what we want.

Junction, intersection n. [du:jəŋga] دُوجنْگا

The order of everything is now correct, but we still have some issues that need to be corrected in Word. Pathway formats the headword and pronunciation both have hyperlinks that are clutter in a printed document, and they also have hard-coded underlines. Section 2.4.2 explains how to get rid of the hyperlinks and the underlines.

There is also another issue from Pathway. The open square bracket for the pronunciation is Doulos SIL coming from the lexemeform-pb1 style, and the closing bracket is Times New Roman in the lexemeform-pa1 style without an actual font specified. As a result, they have a slightly different height. This can be fixed by changing the lexemeform-pb1 style to use Doulos SIL in Word.

There is another problem introduced by Pathway, however, for words that are not immediately following a letter header.

د

[[du:jəŋga] Junction, intersection n. دُوجنْگا دُوج

This is the same entry that worked fine when it was directly after the letter header. But in all other locations, there is now an extra square bracket before the pronunciation, and the pronunciation ends up following the definition. There are two approaches to fixing this.

1. After opening the .doc file in Word, you can carefully highlight the double square brackets, which have a hidden mark between them and copy this into the Find and Replace Find box and place a single open square bracket in the Replace box. When the search and replace are done, the double square bracket is replaced by a single square bracket and the position of the pronunciation goes back to the desired location.
2. This can also be removed in the .doc file produced by Pathway before opening the file in Word. The problem is that Pathway inserted an extra span with the LRE code (U+202A) before the opening square bracket (shown in red here)
 

```
<span class="lexemeform-pb" lang="trw-fonipa">LRE[</span><span class="lexemeform-pb" lang="trw-fonipa" dir="ltr">LRE[</span>
```

 and that did not have a closing PDF code, thus causing the wrong positioning of the pronunciation. This can be fixed with search and replace, or a CC table

Search: `<span class="lexemeform-pb" lang="trw-fonipa">LRE[</span><span class="lexemeform-pb"`  
Replace: `<span class="lexemeform-pb"`

The discussion above illustrates some of the kinds of challenges that will likely be encountered when configuring Bidi dictionaries in FLEx, and some specific challenges when using Word, and further challenges when using Pathway. Any time you add Before and After fields, there is a good possibility you will need to use some of the bidi code points to get the desired result. It's also important to test all of these after importing into Word to make sure they work the same way as they do in FLEx. Here are some examples of bidi code points added to FLEx configuration.

- For sense number: Before: [RLO] After: )][PDF][NBSP] to keep together and to get correct RTL ) glyph.
- One way to reduce the bidi complexities is to start some fields on a new line: Before: \A
- For scientific name: Before: \A[LRE]Sc.[NBSP]Name:· After: [PDF] to keep together and start a new line.
- For Dialect Labels Before: [RLO][ After: ]][PDF]·

---

<sup>i</sup> Macro from [https://wordribbon.tips.net/T000997\\_Preventing\\_Styles\\_from\\_Changing.html](https://wordribbon.tips.net/T000997_Preventing_Styles_from_Changing.html)

<sup>ii</sup> Macro from <https://www.extendoffice.com/documents/word/5423-word-convert-linked-image-to-embedded.html>.

<sup>iii</sup> Macro from <https://www.extendoffice.com/documents/word/5423-word-convert-linked-image-to-embedded.html>

<sup>iv</sup> Macro from [https://word.tips.net/T001337\\_Removing\\_Unused\\_Styles.html](https://word.tips.net/T001337_Removing_Unused_Styles.html)