

***Use PC-PATR with FLEx* User Documentation**

H. Andrew Black

SIL International

andy_black@sil.org

September 26, 2018

Version 0.1.0 Beta

Copyright © 2018 SIL International

Contents

1	Introduction	1
2	Buttons	2
2.1	PC-PATR grammar Browse button	2
2.2	Choose <i>FLEx</i> project button	2
2.3	Help button	3
2.4	Disambiguate button	3
2.5	Parse button	3
3	Parsing a segment	3
4	Disambiguating a text	4
5	Restarting <i>Use PC-PATR with FLEx</i>	4
6	Error messages	4
7	Known problems	5
8	Support	5
A.	The <i>PcPatr Browser</i> tool	5
B.	Another way to edit a PC-PATR grammar file	7

1 Introduction

Use PC-PATR with FLEx is a tool that can use a PC-PATR grammar to disambiguate a text or a portion of a text that has been analyzed in a *FieldWorks Language Explorer* project. You tell *Use PC-PATR with FLEx* the PC-PATR grammar file and the *FLEx* project to use. Then you can choose a text or a portion of that text and ask *Use PC-PATR with FLEx* to try and disambiguate it.

While you can use any PC-PATR grammar file, *Use PC-PATR with FLEx* expects that you will have used the *PAWS* program to create the PC-PATR grammar file. See <https://software.sil.org/paws/> for more on this program.

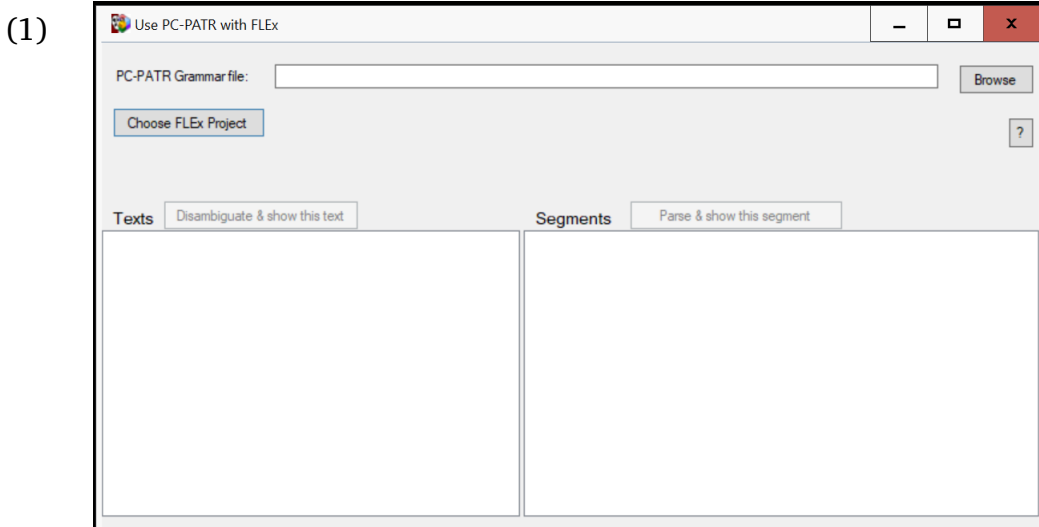
Whenever you first load a *FLE_x* database in *Use PC-PATR with FLE_x*, it will automatically add to your *FLE_x* database the following:

1. a custom field to each sense (called “PCPATR”) and
2. a custom list (called “PCPATR Feature Descriptors”) containing the template names that the PC-PATR grammar generated by the *PAWS* program expects.

The names shown above are always in the English analysis writing system and English is the only writing system containing these names.

Please be aware that you cannot run both *Use PC-PATR with FLE_x* and *FLE_x* at the same time. You will need to close one before you start the other.

Use PC-PATR with FLE_x looks like what is shown in (1).



There are five buttons you can click. Each is discussed in section 2 below.

2 Buttons

You control *Use PC-PATR with FLE_x* by using the various buttons. This section briefly describes them. See sections 3–4 for more on what happens for the segment and text buttons, respectively.

2.1 PC-PATR grammar Browse button

To choose which PC-PATR grammar file to use, click on the Browse button. By convention, PC-PATR grammar files have an extension of “.grm” so this is what the file browser uses.

2.2 Choose *FLE_x* project button

When you click the “Choose FLE_x Project” button, you will see the standard *FLE_x* choose project dialog box. You use it to select the project to use. Once the project

has been loaded (which may take a while, depending on the size of the project), the names of the project's texts (in the default analysis language) will appear in the bottom left pane. By default, the first text will be selected and all of its segments¹ will be shown in the bottom right pane (using the default vernacular language). By default, the first segment will also be selected.

Note that if a given segment does not yet have an analysis for every word in it, the segment will be grayed out. This is because there is no need to try and pass this segment to the PC-PATR grammar. Since at least one word does not have an analysis, *PC-PATR* cannot parse it: there is no parsing information for that word so the segment is guaranteed to fail to parse.

Please note that you cannot run both *FLEX* and *Use PC-PATR with FLEX* on the same project at the same time. You will need to close one before you start the other.

2.3 Help button

The smaller button with a question mark in it is used to get this user documentation file or to show the “About” dialog box.

2.4 Disambiguate button


Above the pane containing the texts is a button labeled “Disambiguate & show this text.” You use this button to try and use the PC-PATR grammar to disambiguate this entire text. See section 4 for more.

2.5 Parse button

Above the pane containing the segments of the selected text is a button labeled “Parse & show this segment.” You use this button to try and use the PC-PATR grammar to parse this particular segment. See section 3 for more.

3 Parsing a segment

When you use the “Parse & show this segment” button for a segment, *Use PC-PATR with FLEX* will pass the parsing information that you entered in *FLEX* to *PC-PATR*, along with the PC-PATR grammar, and process that segment. The result will be shown in a slightly newer version of the *PcPatr Browser* tool (<https://software.sil.org/carla/pc-patr-browser/>). See appendix A for more information on this tool.

You use this tool to see all the parses the *PC-PATR* program produced for this segment using the PC-PATR grammar file. When there are two or more parses, you can show the correct one and then use **Parse / Use This Parse** menu item (or click on the  tool bar button) to have this analysis be recorded in the *FLEX* database for this segment. That is, this segment will be disambiguated, using the information from this parse.


¹By “segment” we mean the portion of the text that *FLEX* treats as a unit. These are often sentences or clauses.

When you are done looking at the results, close the *PcPatr Browser* tool and you will return to *Use PC-PATR with FLEEx*.

4 Disambiguating a text

When you use the “Disambiguate & show this text” button, *Use PC-PATR with FLEEx* will pass the parsing information for all segments in the text to *PC-PATR*, along with the PC-PATR grammar, and process all those segments. Since all the segments are processed, this will probably take a while. So please be patient. The result will be shown in the slightly newer version of the *PcPatr Browser* tool. See appendix A for more information on this tool.

Unlike with parsing a segment (see section 3), the result will have as many “sentences” as there are segments in the text.

By default, *Use PC-PATR with FLEEx* will disambiguate a segment in the *FLEEx* database whenever the result of the parsing uniquely identifies the analysis of each word in the segment. For those cases where there are still ambiguities, you can use the *PcPatr Browser* tool to select a given parse to use. Like with parsing a segment, you show the correct parse and then use **Parse / Use This Parse** menu item (or click on the  tool bar button) to have this analysis be recorded in the *FLEEx* database for this segment. That is, this segment will be disambiguated, using the information from this parse.

When you are done looking at the results, close the *PcPatr Browser* tool and you will return to *Use PC-PATR with FLEEx*.

5 Restarting *Use PC-PATR with FLEEx*

Whenever you exit and restart *Use PC-PATR with FLEEx*, it will do the following:

1. remember which PC-PATR grammar file you last chose;
2. remember which *FLEEx* project you last used and open it;
3. remember which text in that project you last selected; and
4. remember which segment in that text you last selected.

As a result, if the last *FLEEx* project you used is large, it may well take a while before *Use PC-PATR with FLEEx* finishes starting.

6 Error messages

In certain situations, *Use PC-PATR with FLEEx* will issue an error message.

Example (2) lists the errors *Use PC-PATR with FLEEx* reports along with a brief description of what the error might mean.

(2)	Error	Meaning
	Loading failed. The FieldWorks project is currently open in another application. Close the application and try to run this command again.	Most likely, <i>FLEX</i> is open with the same project. Please close <i>FLEX</i> .

If you get an error message not in the list above, please report it. See section 8.

7 Known problems

The following items are known to be less than desirable with this version of *Use PC-PATR with FLEX*:

- When a word in *FLEX* does not have the “Word Cat” field filled out for a word in the interlinear, *Use PC-PATR with FLEX* will try and guess what the category of the word should be. It does this as follows:
 1. If there is only one morpheme, use its category.
 2. If there are two or more morphemes and only one stem/root, find the stem and use its category. (This could be incorrect if one of the affixes is derivational and changes the category.)
 3. If there are two or more stems, use the category of the left-most one. (This could be incorrect if some other stem is the correct category to use or if the stem compounding results in a different category or if one of the affixes is derivational and changes the category.)
- *Use PC-PATR with FLEX* assumes that the comment character in the PC-PATR grammar file is a vertical bar |. If you use something else, the parser will fail every segment.

8 Support

If you have any questions with *Use PC-PATR with FLEX* or find bugs in it, please send an email to andy_black@sil.org.

A. The *PcPatr Browser* tool

This appendix briefly covers the various parts of the *PcPatr Browser* tool.

Example (3) is a screen shot of what it looks like for one segment using one PC-PATR grammar.

(3)

The screenshot shows the PC-PatRBrower application interface. The title bar indicates the file path: C:\Users\ANDYBL-1\AppData\Local\Temp\Invoker.and (Default Language). The menu bar includes File, View, Sentence, Parse, Language, and Help. The toolbar contains various icons for file operations and parsing.

Sentence 1 of 1:

we	want	to	get	married	and	be	happy
1PL-nom	want	to	get-cop	marry-PART	and	be	happy
pro	v	aux	v	v	conj	v	adj
first nominative plural	sentential_-finiteIP_or_transitive	-finite	copular	participle	plural	copular	

Parse 1 of 2:

```

      S
      |
      IP
     /  \
  DP+    I'
  /  \   /  \
D'+   V    VP
|      |   /  \
Pron   V    IP
|       |   /  \
we     want  Aux  VP
1PL-nom want
  
```

Feature structure of the selected node (IP):

When you click on a node in the tree shown in the panel to the left, any feature structure associated with that node will show up here. (Note: if you do not have PC-PATR include feature structures in the output, then no features will ever show here.)

PC-PATR Grammar file for PAWS Starter Kit

DO NOT EDIT THIS FILE UNTIL YOU HAVE COMPLETED THE ENTIRE STARTER KIT!
(Otherwise your changes will get over-written the next time you save.)

PC-PATR Grammar file for PAWS Starter Kit
C Black July 2012

Phrase structure rules

Let ablative be <head case> = ablative

Sentence 1 of 1 Parse 1 of 2

There are four panes:

1. the segment in an interlinear form (at the top);
2. the syntactic tree, if any (left, middle);
3. the feature structure of the selected node in the tree (right, middle); and
4. the PC-PATR grammar file used (bottom)

When a node in the tree is selected, the bottom pane will show the rule in the PC-PATR grammar file that produced that node. In addition, the features for that selected node will appear in the right, middle pane. For example, when clicking on the IP node, it looked like example (4).

(4)

The screenshot shows the PcPatrBrowser application window. The title bar indicates the file path: C:\Users\ANDYBL~1\AppData\Local\Temp\Invoker.and (Default Language). The menu bar includes File, View, Sentence, Parse, Language, and Help. The toolbar contains various navigation and editing icons.

Sentence 1 of 1

we	want	to	get	married	and	be	happy
we	want	to	get	married	and	be	happy
1PL-nom	v	aux	v	marry-PART	conj	v	happy
pro	sentential_finiteIP_or_transitive	-finite	copular	participle	plural	copular	adj
first nominative plural							

Parse 1 of 2

The parse tree shows a hierarchical structure starting with S, branching into IP and I'. The IP branches into DP+ and VP. DP+ branches into D'+, which further branches into Pron (we, 1PL-nom, pro). The VP branches into V (want, v, want) and another IP. This second IP branches into Aux (to, aux, -finite) and a final VP (get, v, get-cop; married, v, marry-PART; and, conj, and; be, v, be; happy, adj, happy).

Feature structure for the final IP:

```
[cat: IP
  head: $10 [adjoinedVP: -
             adverbialVP: -
             embedded: $7 [cat: IP
                           head: $6 [adjoinedVP: -
                                     adverbialVP: -
                                     embedded: [cat: none]
                                     fronted: none
                                     infl: [finite: -
                                           polarity: positive
                                           valence: active
                                           voice: none ]
                                     subject: head: faer: 11 ]
                           ]
             ]
```

Rule definition:

```
rule (IP option 2a - subject initial, required, active only)
IP = DP I'
  <IP head> = <I' head>
  <IP head type root> = <DP head type root>
  <IP head subject pro-drop> = -
  <I' head subject> = <DP>
  <I' head infl valence> = active
  <DP head type focusmarked> = - | + in focus position
  <DP head type prefix focusmarked> = - | + in focus position
  <DP head type suffix focusmarked> = - | + in focus position
  <I' head type existential> = -
  <I' head type speech_PP> = +
  <I' head type speech_DP> = +
  <I' head type copular> = +
  <I' head type perception> = +
```

When there is more than one segment, you can use the **Sentence** menu item options or the first set of arrow buttons on the tool bar to navigate from one sentence to another. (Please note that with *Use PC-PATR with FLEx*, “sentence” is the same as “segment.”)

When there is more than one parse for a given segment, you can use the **Parse** menu item options or the second set of arrow buttons on the tool bar to navigate from one parse to another.

B. Another way to edit a PC-PATR grammar file

One challenge with editing a PC-PATR grammar file is that it is a plain text file. If one wants to try and debug a grammar, one has to comment off a large amount of lines. It can get to be tricky.

Some tentative work has been done on a tool that will take a PC-PATR grammar file and convert it to a form that can be used in a tool like the *XMLmind XML Editor*. Some advantages of this are that one can:

1. Quickly enable/disable given rules
2. Quickly enable/disable given templates

A major disadvantage is that once converted to this form, large PC-PATR grammar files are quite slow to load into the *XMLmind XML Editor*.

If you are interested in trying this, please contact andy_black@sil.org.