P_Lex v.3.0: The Manual
Paul Meara

1.   P_Lex was designed as a way of assessing the vocabulary used by L2 learners when they produce texts. The program divides a text up into ten-word segments and counts the number of "difficult" or "unusual" words that appear in each segment. The score it generates is the mean number of difficult words in each segment.

P_Lex works with short texts in English and is mainly intended to process short texts generated by L2 learners. As such, it works best with texts that are not longer than 300 words. The program will handle texts that are longer than this, but the displays are designed for shorter texts.

2. You will need to do some **editing work** on the texts that you want to process using P_Lex. The main problem that you will have to deal with concerns errors. There are two possibilities here. You can leave the errors intact or you can correct them. Since P_Lex uses an inbuilt dictionary that will not recognise incorrectly spelled words, it is probably best to correct errors of this sort.

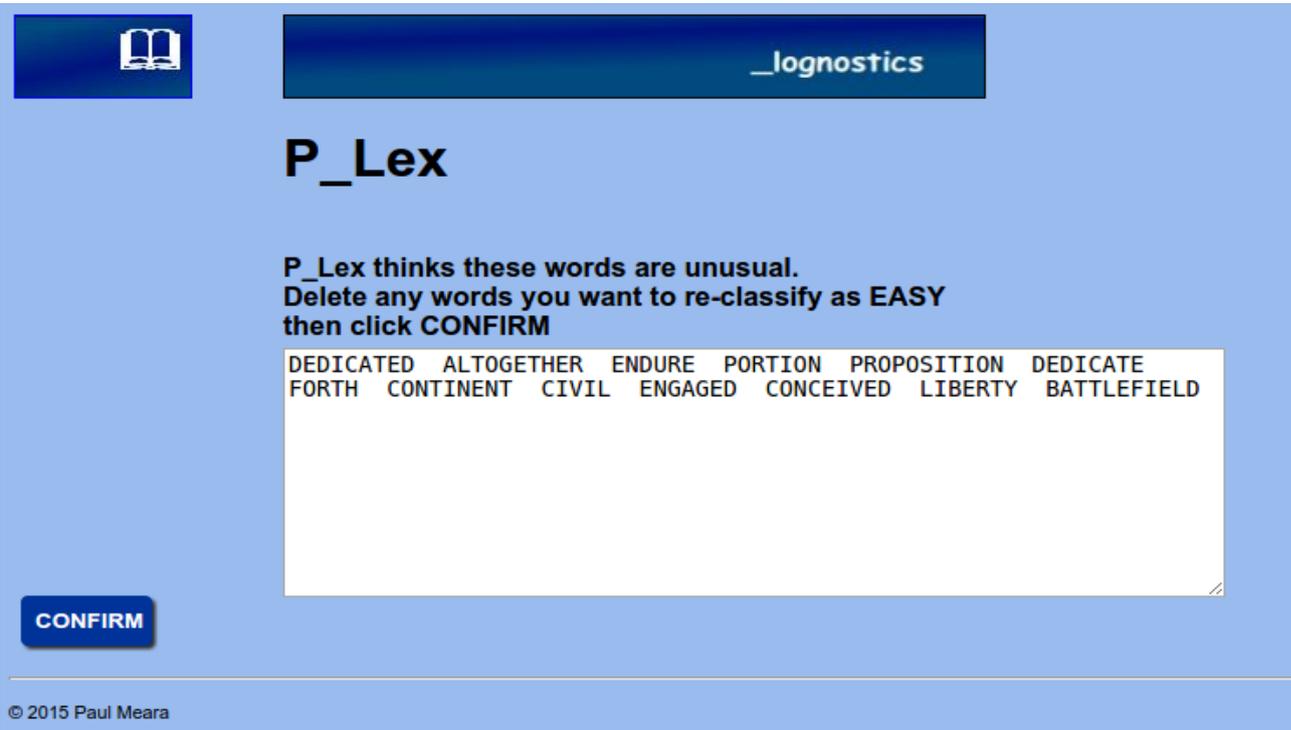3. The P_Lex **work space** is shown below:

4. **Load the tex**t that you want to work with into the P_Lex workspace. You can do this by typing your text directly, or by copying a pre-prepared text into the workspace. Add an identifying label for this text in the box underneath the main workspace.

5. P_Lex will ignore any punctuation marks or other symbols in the text that are included in the list underneath the main workspace. You can change these symbols, or add new ones if you need to.

6.   Click the **Submit** button to analyse your text. P_Lex will generate a preliminary report which includes any words that it does not recognise. This report will look like the figure below. P_Lex has an
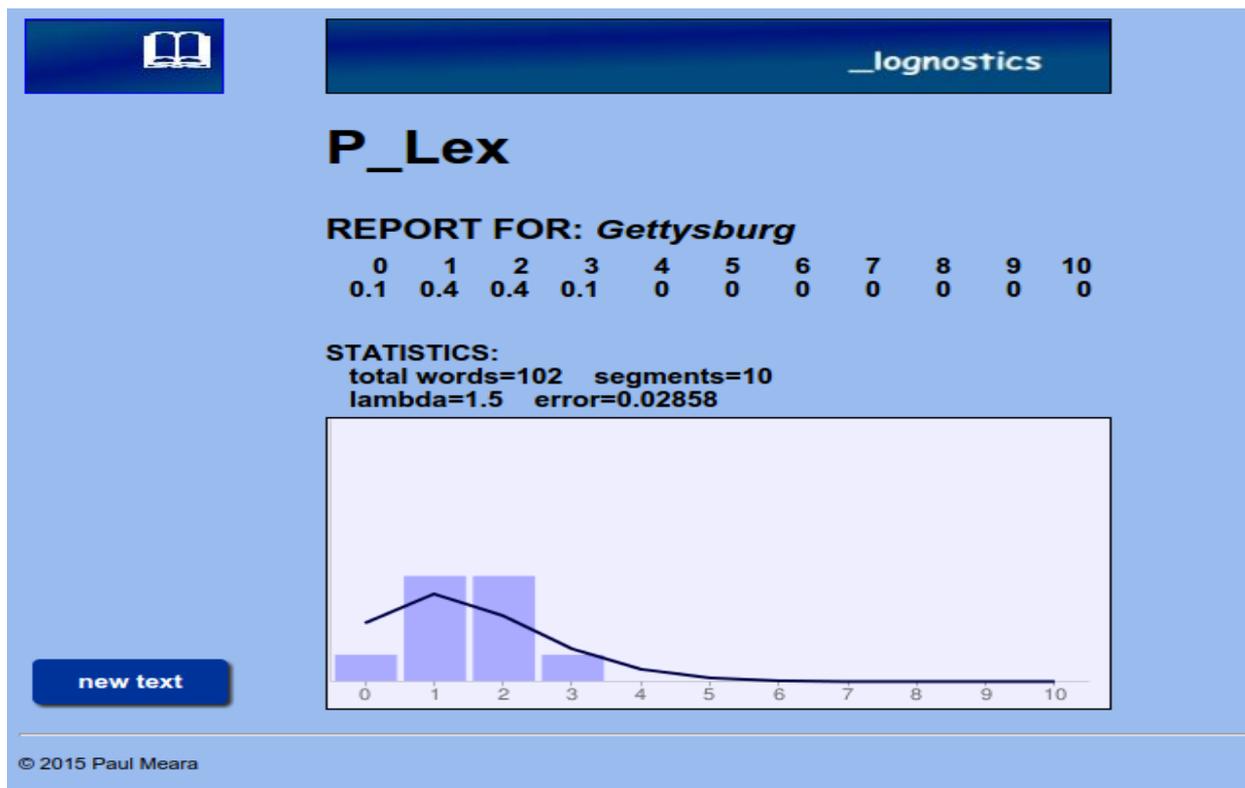
in-built dictionary file that includes a number of highly frequent words, and it treats any word that does not appear in this internal dictionary list as "unknown". Generally, these words will be low frequency words. However, you may not agree with the classification that P_Lex makes. For example, if you are dealing with a group of Icelandic learners of English who are describing their home town, then the texts they generate are very likely to include *Reykjavik*. P_Lex would report this item as an unusual word, but you would probably want to override this classification. P_Lex will also report misspelled or invented words as unknown. Again, you will probably want to override this classification.

You can override P_Lex in two ways when you want to take out words from this 'unusual words' list. The first way is to go back to the original input screen and edit your text in the workspace. You might want to use this method to correct spelling errors, for example. Click **Submit** again when you are satisfied that your editing is complete, and P_Lex will reanalyse your text. The second way is to delete items directly from the unknown words report. P_Lex will add any deleted words to its dictionary file and treat them as easy words when it computes the statistics for your text.

7. When you are happy that the words identified by P_Lex are indeed unusual words, and not errors or misclassified items, then click the **Confirm** button to proceed. P_Lex generates a report that looks like this:

_lognostics

P_Lex

REPORT FOR: *Gettysburg*

0	1	2	3	4	5	6	7	8	9	10
0.1	0.4	0.4	0.1	0	0	0	0	0	0	0

STATISTICS:
total words=102   segments=10
lambda=1.5   error=0.02858

new text

The report tells you the name of the text it has analysed, how many 10-word segments have been processed by the program, and what proportion of the ten-word segments contain N difficult words. It also reports a value (lambda) which summarises the data presented in the table. This is the value that you would normally use to summarise a P_Lex analysis. For the complete explanation on how this value is calculated, please read the article in the 'Background reading section' (also reproduced in Meara & Miralpeix 2017).

The solid line superimposed on the data tells you how well the P_Lex model fits the actual data. In the example above, the data and the model fit moderately well, and the error figure is correspondingly low. However, this result needs to be treated with caution when the text analysed is very short.

8. Click on the **New text** button if you want to analyse more texts.

Background reading:

Meara, P.M. and Bell, H. (2001) P_Lex: A simple and effective way of describing the lexical characteristics of short L2 texts. *Prospect* 16 (3), 5-19.

For a more detailed discussion of *P_Lex*, see **PM Meara and I Miralpeix (2017)** *Tools for Researching Vocabulary*. Bristol: Multilingual Matters.