

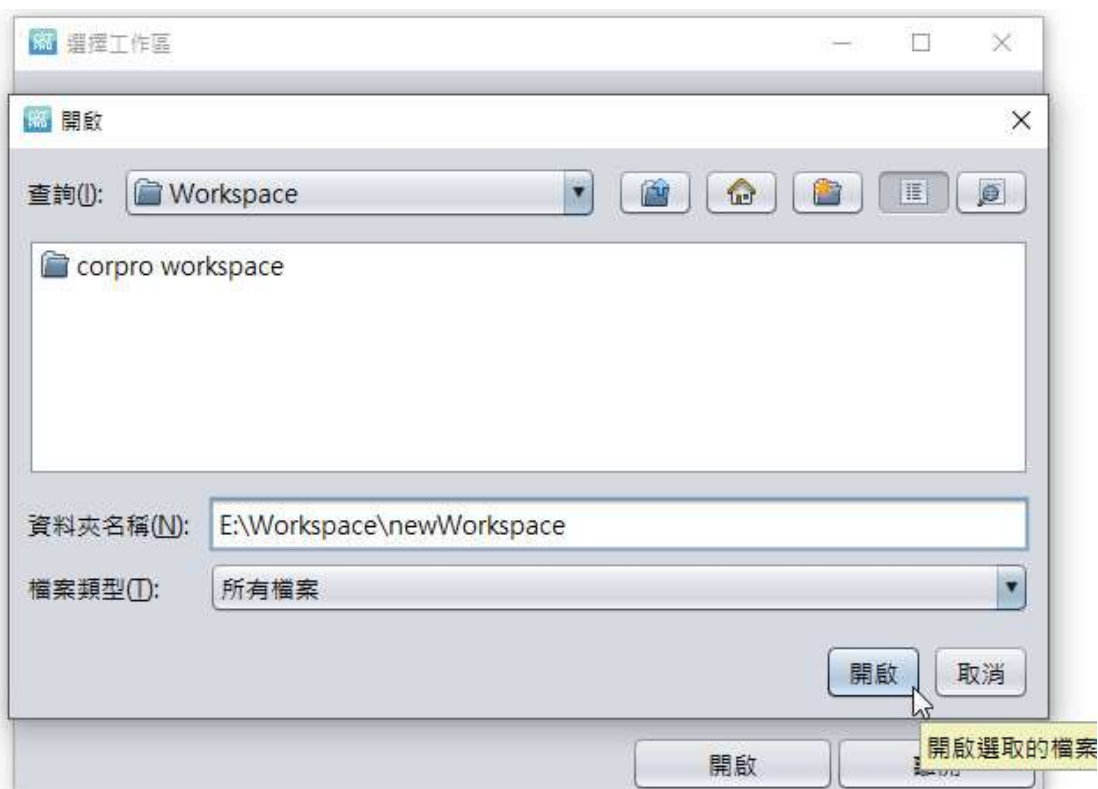
一、開始使用 CORPRO

- ❖ 執行 CORPRO.bat 檔
- ❖ 出現 CORPRO 開始畫面

功能：

1. 建立新的工作區 (按「瀏覽」先翻找到要建工作區的資料夾，再直接在「資料夾名稱」處加上「你的工作區名字」)

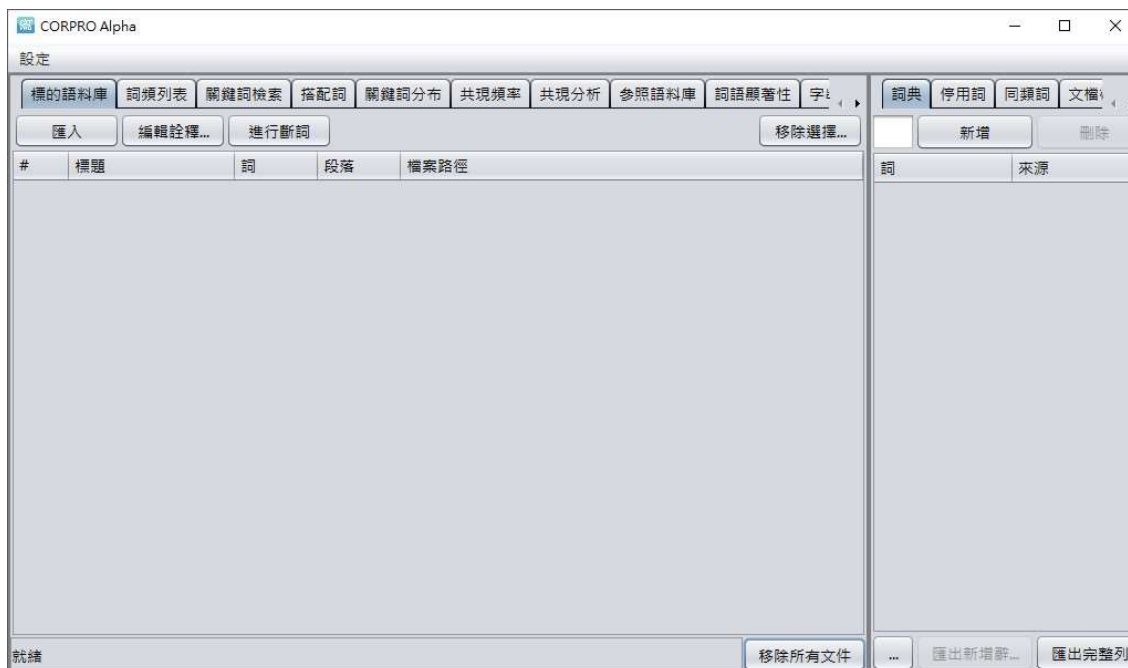
(範例: 想建立一個新的工作區叫作 newWorkspace，按「瀏覽」翻找到要建工作區的資料夾 E:\Workspace，再在後面打字加上 newWorkspace 成為 E:\Workspace\newWorkspace 即可)



2. 選擇之前的工作區 (demo)



❖ 出現 CORPRO 主畫面



二、匯入語料庫 (標的語料庫或參照語料庫)

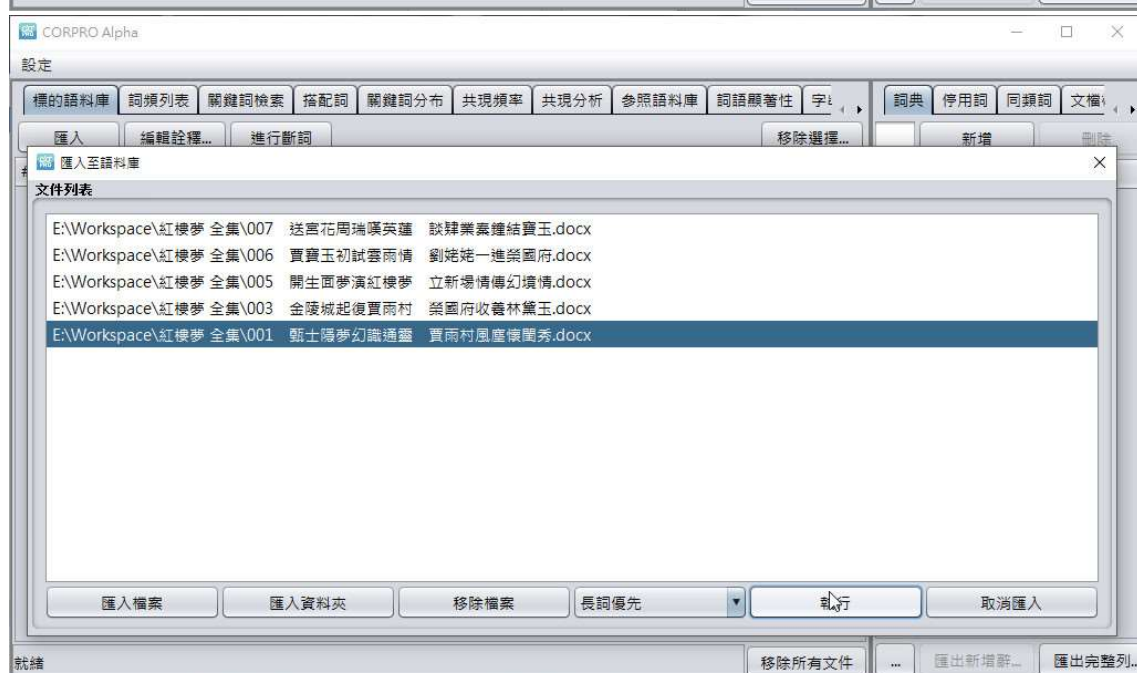
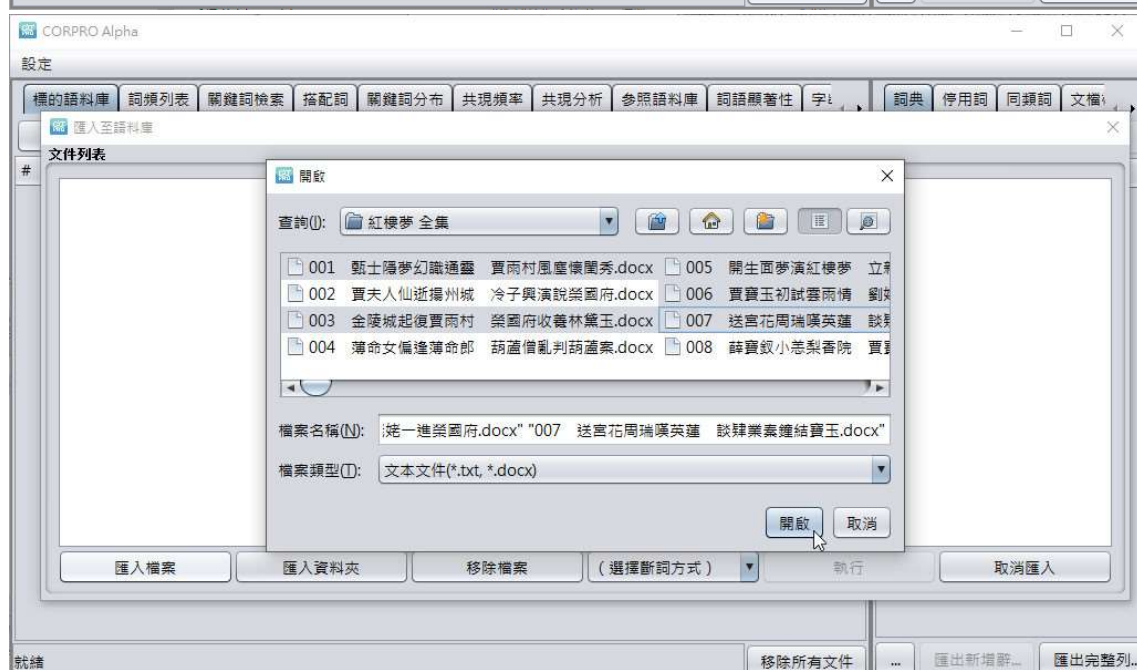
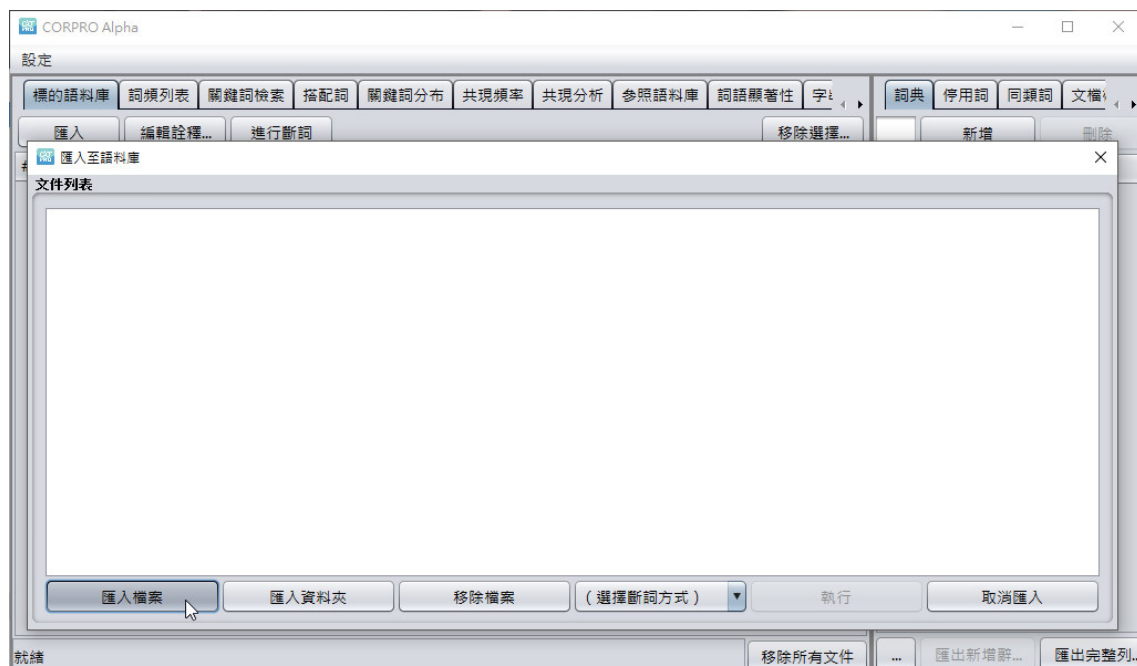
- ❖ 按「匯入」按鈕
- ❖ 出現匯入語料庫畫面

功能：

1. 匯入檔案

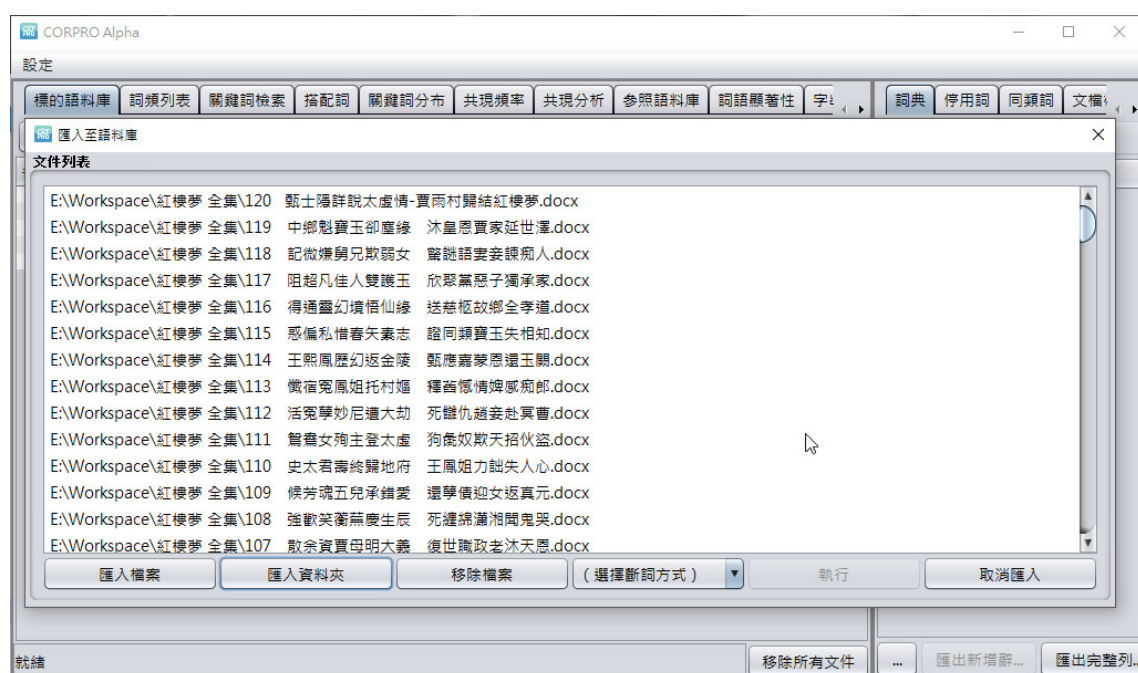
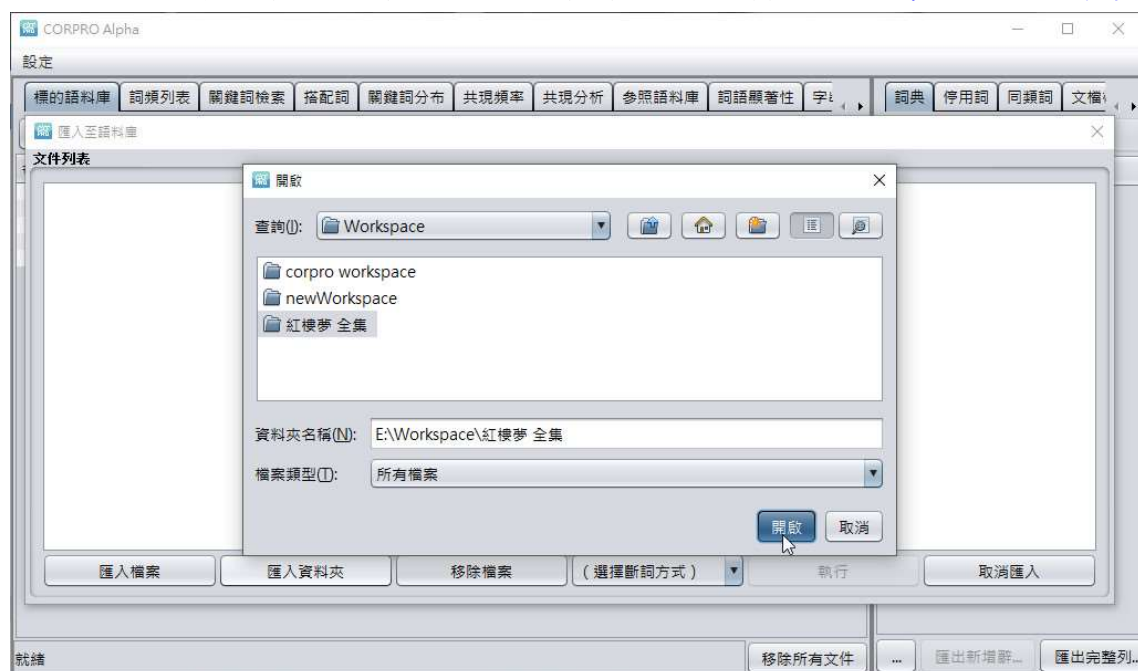
出現檔案選擇對話視窗，單選、連續多選、不連續多選都沒問題

檔案類型可為純文字檔或是 MS Word 檔 (demo: 選紅樓夢第一、三、五~七回)



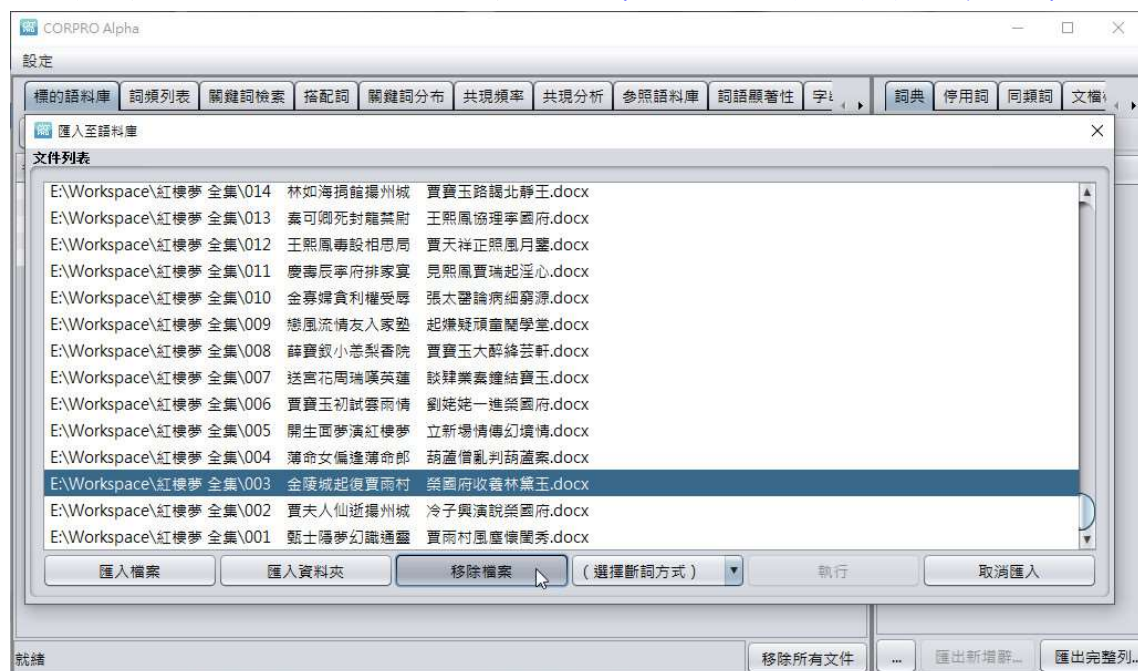
2. 匯入目錄

出現目錄選擇對話視窗，所選的目錄底下所有檔案都會全部匯入(demo: 選紅樓夢資料夾)



3. 移除檔案

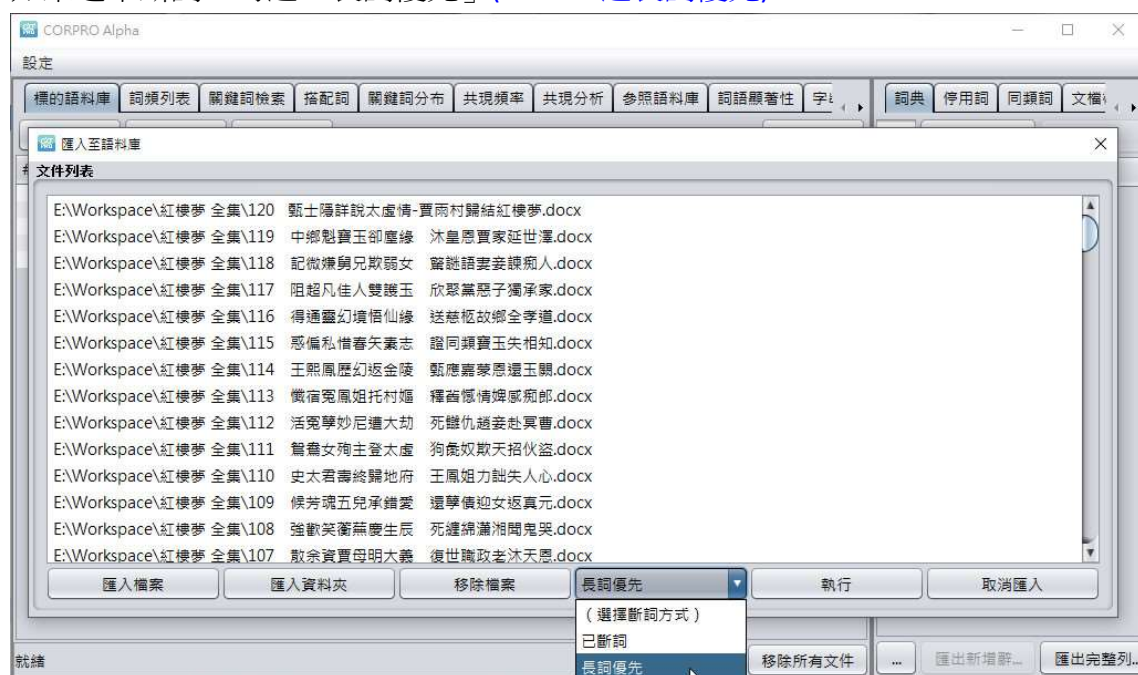
前兩步所選的檔案在未匯入前可先移出列表(demo: 上一步之後移除第 3 回)



4. 選擇斷詞方式

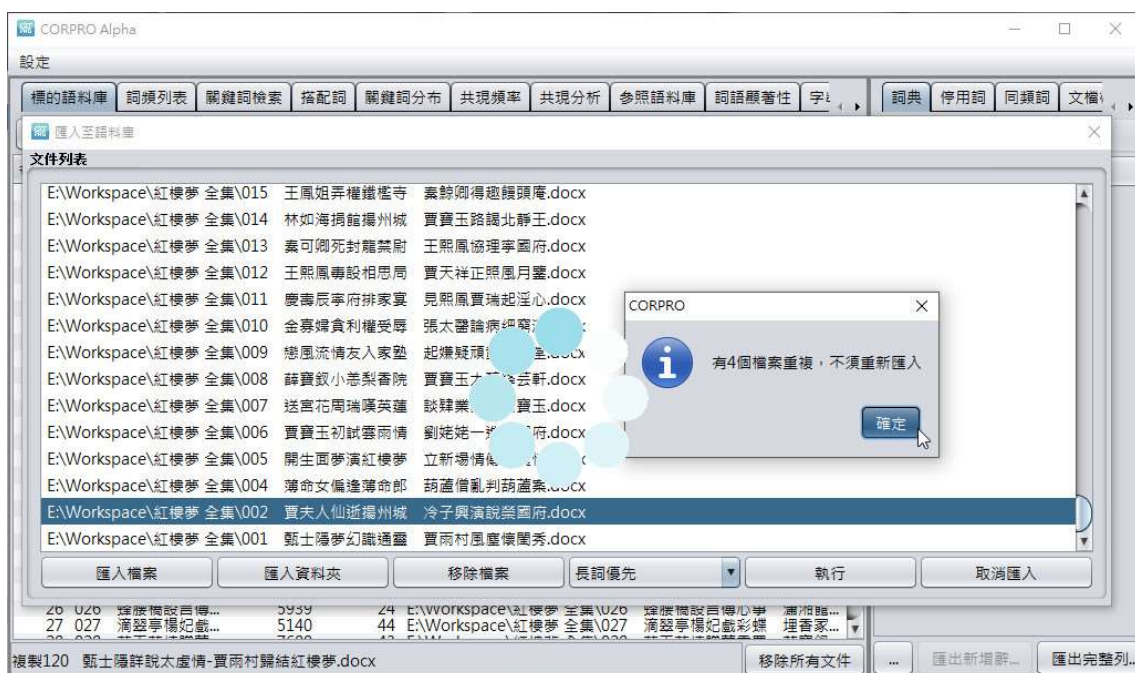
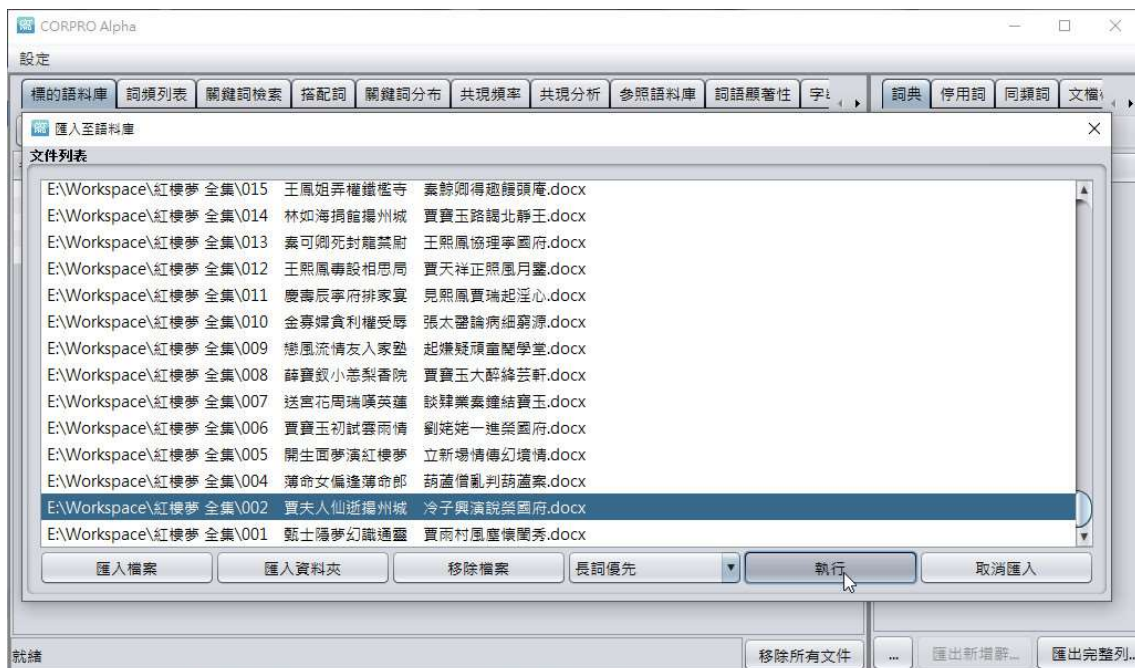
如果要匯入的檔案已經斷過詞，就選「已斷詞」

如果還未斷詞，可選「長詞優先」(demo: 選長詞優先)



5. 執行匯入

按「執行」開始匯入後，在列表中、但早已匯入語料庫的文件檔會先跳「不會匯入」的通知。列表中所有其他文件檔都會被斷詞（若尚未斷詞）並加入語料庫中。（demo: 跳出 1,5,6,7 已存在的訊息、匯入完成後回到標的語料庫的畫面）

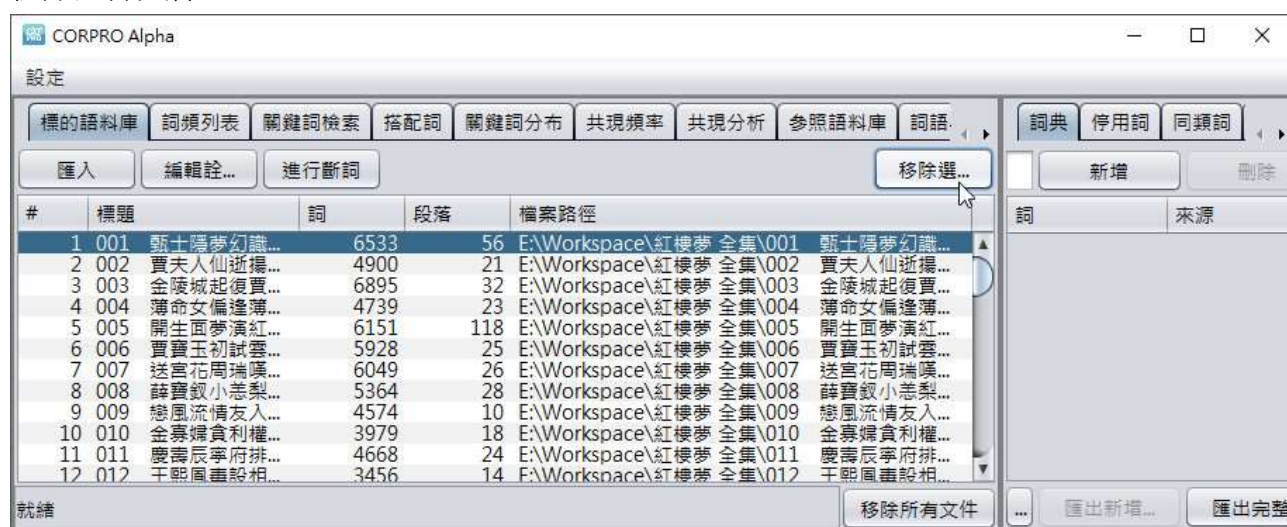


❖ 回到標的語料庫

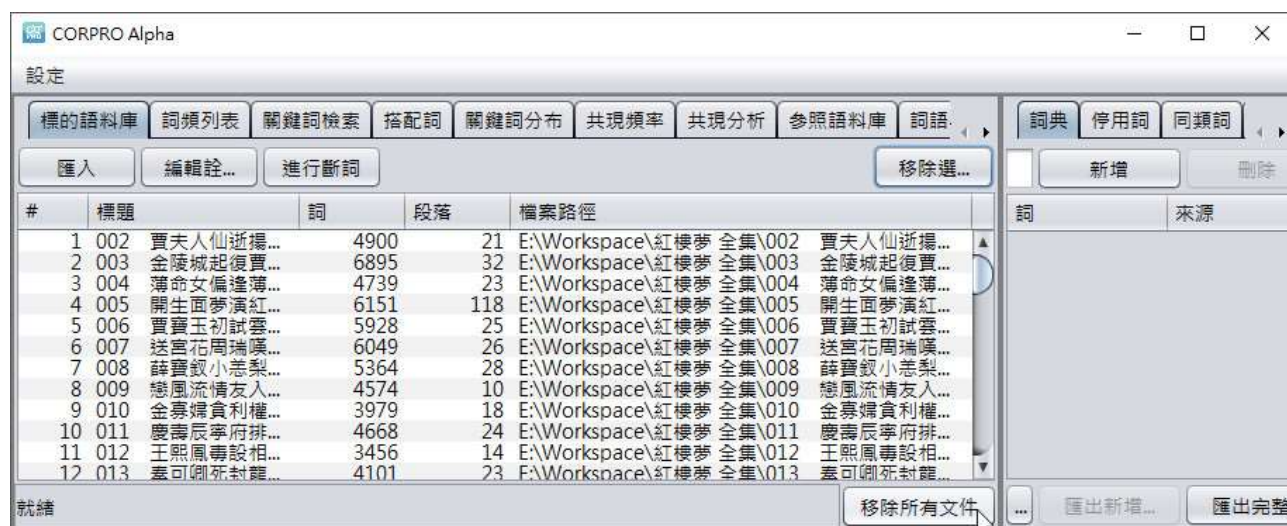
三、標的語料庫

功能：

移除選擇文件



移除所有文件



四、文檔檢視

功能

1. 點兩下開啟



2. 搜尋詞

The screenshot shows the CORPRO Alpha software interface. The main window displays a list of documents with columns for #, 標題 (Title), 詞 (Word), 段落 (Paragraph), and 檔案路徑 (File Path). The search results pane on the right shows the text of the selected document, with the character '寶玉' highlighted in blue. The search results pane also includes a search bar and a list of search results.

3. 跳下一個

The screenshot shows the CORPRO Alpha software interface. The '跳下一個' (Jump to Next) button is highlighted in the toolbar. The main window displays a list of documents, and the search results pane on the right shows the text of the selected document.

4. 清空

The screenshot shows the CORPRO Alpha software interface. The '清空' (Clear) button is highlighted in the toolbar. The main window displays a list of documents, and the search results pane on the right shows the text of the selected document.

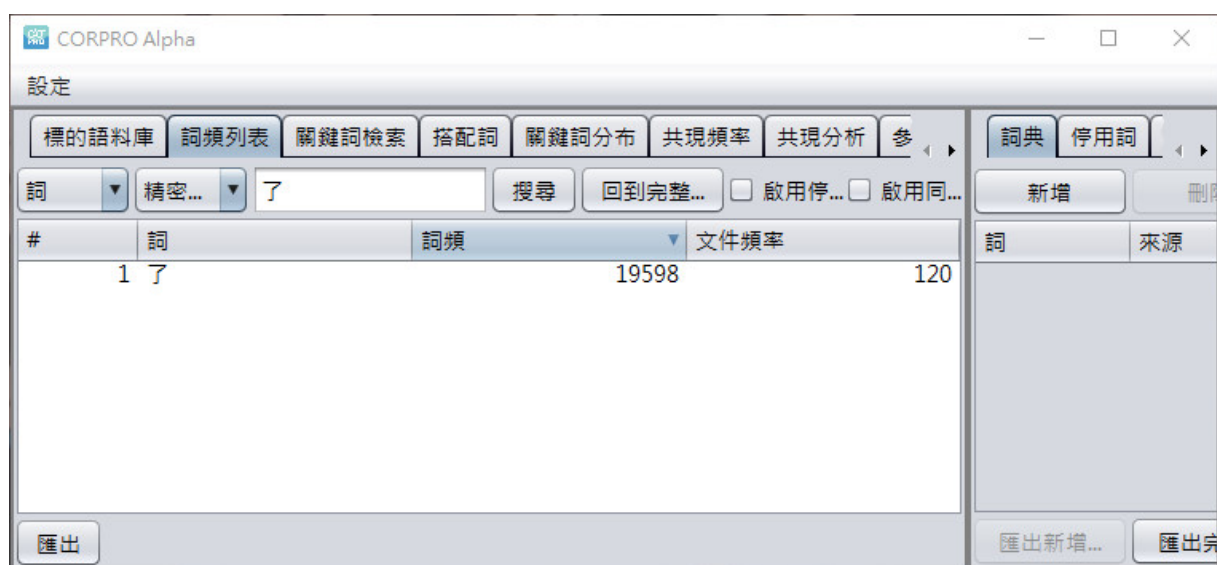
五、詞頻列表

功能

1. 各欄位可排序



2. 搜尋「詞」+「精密比對」



3. 搜尋「詞」+「子字串比對」



4. 搜尋「詞」+「開頭」



設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現表 詞典

詞 開... 了 搜尋 回... ☐ 啟用... ☐ 啟用...

#	詞	詞頻	文件頻率
1	了	19598	120
2	了去	157	80
3	了不得	74	48
4	了事	46	31
5	了得	38	27
6	了結	16	11
7	了無	6	6
8	了悟	2	1
9	了解	2	2
10	了卻	1	1

匯出

5. 搜尋「詞」+「結尾」



設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現表 詞典

詞 結... 了 搜尋 回... ☐ 啟用... ☐ 啟用...

#	詞	詞頻	文件頻率
1	了	19598	120
2	罷了	295	100
3	好了	231	88
4	就是了	188	90
5	得了	175	86
6	未了	69	42
7	除了	45	29
8	不了	38	31
9	不見了	32	19
10	認了	20	15
11	不得了	19	19

匯出

6. 搜尋「詞頻/文件頻率」+「>/=/<」

CORPRO Alpha

設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現

詞... > 1445 搜尋 回... ☐ 啟用... ☐ 啟用...

#	詞	詞頻	文件頻率
1	,	59310	120
2	。	29406	120
3	了	19598	120
4	的	14447	120
5	」	11773	117
6	：	11749	120
7	「	11192	117
8	我	7750	120
9	他	6597	120
10	說	6120	120
11	你	6080	120

匯出

CORPRO Alpha

設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現

詞... = 1445 搜尋 回... ☐ 啟用... ☐ 啟用...

#	詞	詞頻	文件頻率
1	夫人	1445	109

匯出

CORPRO Alpha

設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現

詞... < 1445 搜尋 回... ☐ 啟用... ☐ 啟用...

#	詞	詞頻	文件頻率
1	王	1424	109
2	心	1351	120
3	問	1323	120
4	忙	1318	115
5	回	1269	119
6	沒	1238	117
7	才	1221	118
8	中	1218	120
9	我們	1200	119
10	吃	1196	112
11	倒	1189	120

匯出

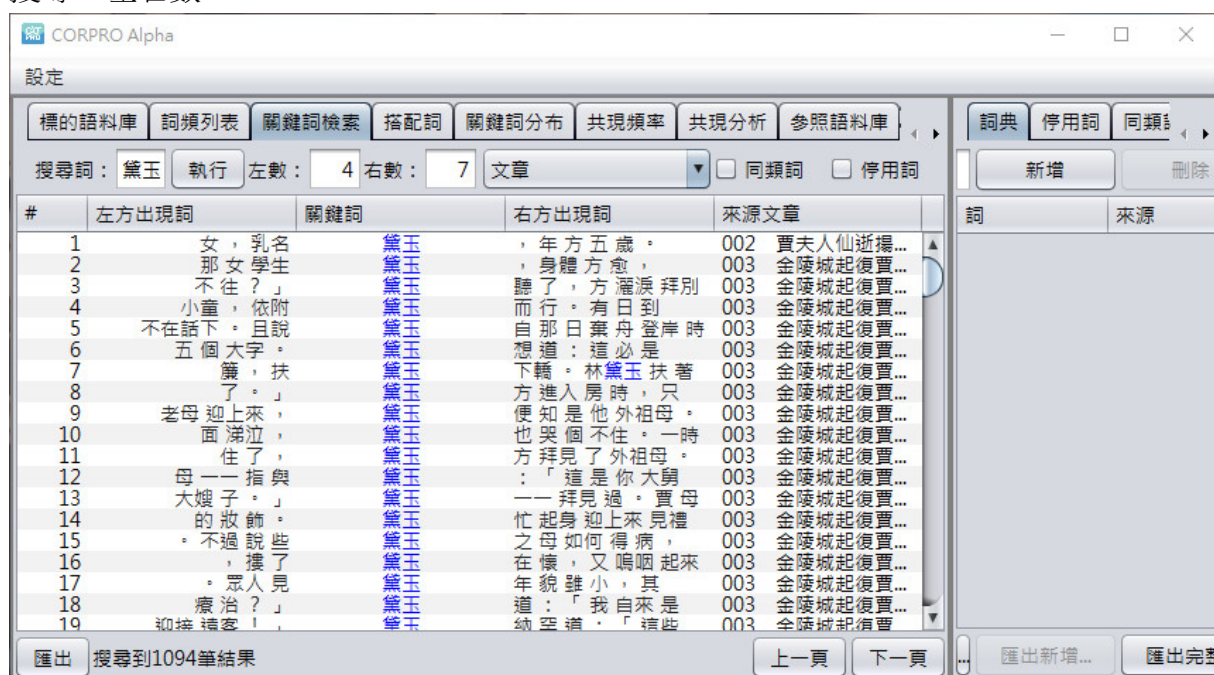
7. 回到完整列表



六、關鍵詞檢索

功能

1. 搜尋、左右數



2. 各種邊界

CORPRO Alpha

設定

標的語料庫 詞類列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫

搜尋詞：黛玉 執行 左數：4 右數：7 段落 ☐ 同類詞 ☐ 停用詞

新增 刪除

#	左方出現詞	關鍵詞	右方出現詞	來源文章
1	女，乳名	黛玉	，年方五歲。	002 賈夫人仙逝揚...
2	那女學生	黛玉	，身體方愈，	003 金陵城起復賈...
3	」	黛玉	聽了，方瀟淚拜別	003 金陵城起復賈...
4	小童，依附	黛玉	而行。	003 金陵城起復賈...
5	且說	黛玉	自那日棄舟登岸時	003 金陵城起復賈...
6		黛玉	想道：這必是	003 金陵城起復賈...
7	簾，扶	黛玉	下轎。	003 金陵城起復賈...
8		黛玉	方進入房時，只	003 金陵城起復賈...
9	老母迎上來，	黛玉	便知是他外祖母。	003 金陵城起復賈...
10	面涕泣，	黛玉	也哭個不住。	003 金陵城起復賈...
11	住了，	黛玉	方拜見了外祖母。	003 金陵城起復賈...
12	母一一指與	黛玉	：「這是你大舅	003 金陵城起復賈...
13	」	黛玉	一一拜見過。	003 金陵城起復賈...
14		黛玉	忙起身迎上來見禮	003 金陵城起復賈...
15	不過說些	黛玉	之母如何得病，	003 金陵城起復賈...
16	攪了	黛玉	在懷，又嗚咽起來	003 金陵城起復賈...
17	眾人見	黛玉	年貌雖小，其	003 金陵城起復賈...
18	」	黛玉	道：「我自來是	003 金陵城起復賈...
19	，	黛玉	幼空道：「這些	003 金陵城起復賈...

匯出 搜尋到1094筆結果 上一頁 下一頁 匯出新增... 匯出完整

CORPRO Alpha

設定

標的語料庫 詞類列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫

搜尋詞：黛玉 執行 左數：4 右數：7 句子 ☐ 同類詞 ☐ 停用詞

新增 刪除

#	左方出現詞	關鍵詞	右方出現詞	來源文章
1	女，乳名	黛玉	，年方五歲。	002 賈夫人仙逝揚...
2	那女學生	黛玉	，身體方愈，	003 金陵城起復賈...
3	」	黛玉	聽了，方瀟淚拜別	003 金陵城起復賈...
4	小童，依附	黛玉	而行。	003 金陵城起復賈...
5	且說	黛玉	自那日棄舟登岸時	003 金陵城起復賈...
6		黛玉	想道：這必是	003 金陵城起復賈...
7	簾，扶	黛玉	下轎。	003 金陵城起復賈...
8		黛玉	方進入房時，只	003 金陵城起復賈...
9	老母迎上來，	黛玉	便知是他外祖母。	003 金陵城起復賈...
10	面涕泣，	黛玉	也哭個不住。	003 金陵城起復賈...
11	住了，	黛玉	方拜見了外祖母。	003 金陵城起復賈...
12	母一一指與	黛玉	：「這是你大舅	003 金陵城起復賈...
13	」	黛玉	一一拜見過。	003 金陵城起復賈...
14		黛玉	忙起身迎上來見禮	003 金陵城起復賈...
15	不過說些	黛玉	之母如何得病，	003 金陵城起復賈...
16	攪了	黛玉	在懷，又嗚咽起來	003 金陵城起復賈...
17	眾人見	黛玉	年貌雖小，其	003 金陵城起復賈...
18	」	黛玉	道：「我自來是	003 金陵城起復賈...
19	，	黛玉	幼空道：「這些	003 金陵城起復賈...

匯出 搜尋到1094筆結果 上一頁 下一頁 匯出新增... 匯出完整

CORPRO Alpha

設定

標的語料庫 詞類列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫

搜尋詞：黛玉 執行 左數：4 右數：7 標點 ☐ 同類詞 ☐ 停用詞

新增 刪除

#	左方出現詞	關鍵詞	右方出現詞	來源文章
1	乳名	黛玉		002 賈夫人仙逝揚...
2	那女學生	黛玉		003 金陵城起復賈...
3	」	黛玉	聽了	003 金陵城起復賈...
4	依附	黛玉	而行	003 金陵城起復賈...
5	且說	黛玉	自那日棄舟登岸時	003 金陵城起復賈...
6		黛玉	想道	003 金陵城起復賈...
7	扶	黛玉	下轎	003 金陵城起復賈...
8		黛玉	方進入房時	003 金陵城起復賈...
9		黛玉	便知是他外祖母	003 金陵城起復賈...
10		黛玉	也哭個不住	003 金陵城起復賈...
11		黛玉	方拜見了外祖母	003 金陵城起復賈...
12	母一一指與	黛玉		003 金陵城起復賈...
13	」	黛玉	一一拜見過	003 金陵城起復賈...
14		黛玉	忙起身迎上來見禮	003 金陵城起復賈...
15	不過說些	黛玉	之母如何得病	003 金陵城起復賈...
16	攪了	黛玉	在懷	003 金陵城起復賈...
17	眾人見	黛玉	年貌雖小	003 金陵城起復賈...
18	」	黛玉	道	003 金陵城起復賈...
19	，	黛玉	幼空道	003 金陵城起復賈...

匯出 搜尋到1094筆結果 上一頁 下一頁 匯出新增... 匯出完整

3. 點兩下檢視文件

CORPRO Alpha

設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著

搜尋詞：黛玉 執行 左數：4 右數：7 標點

同類詞 停用詞

#	左方出現詞	關鍵詞	右方出現詞	來源文章
1	乳名	黛玉		002 賈夫人仙逝揚州...
2	那女學生	黛玉		003 金陵城起復賈雨...
3	」	黛玉	聽了	003 金陵城起復賈雨...
4	依附	黛玉	而行	003 金陵城起復賈雨...
5	且說	黛玉	自那日棄舟登岸時	003 金陵城起復賈雨...
6		黛玉	想道	003 金陵城起復賈雨...
7	扶	黛玉	下轎	003 金陵城起復賈雨...
8		黛玉	方進入房時	003 金陵城起復賈雨...
9		黛玉	便知是他外祖母	003 金陵城起復賈雨...
10		黛玉	也哭個不住	003 金陵城起復賈雨...
11		黛玉	方拜見了外祖母	003 金陵城起復賈雨...
12	母一一指與	黛玉		003 金陵城起復賈雨...
13	」	黛玉	一一拜見過	003 金陵城起復賈雨...
14		黛玉	忙起身迎上來見禮	003 金陵城起復賈雨...
15	不過說了些	黛玉	之母如何得病	003 金陵城起復賈雨...
16	攪了	黛玉	在懷	003 金陵城起復賈雨...
17	眾人見	黛玉	年貌雖小	003 金陵城起復賈雨...
18	」	黛玉	道	003 金陵城起復賈雨...
19	」	黛玉	納罕道	003 金陵城起復賈雨...
20		黛玉	連忙起身接見	003 金陵城起復賈雨...

匯出 搜尋到1094筆結果

上一頁 下一頁

002 賈夫人仙逝揚州城 冷子興演說榮國府.docx

詩云

一局輸贏料不真，香燭茶盞尚逡巡。

4. 查詢兩個以上的詞、用空白隔開表示 OR

CORPRO Alpha E:\Workspace\test\CORPRO Workspace\

設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性 字串頻

搜尋詞：黛玉 夫人 執行 左數：2 右數：2 文章

同類詞 停用詞

#	左方出現詞	關鍵詞	右方出現詞	來源文章
1086	便將	黛玉	臨死的話	109 候芳魂五兒承錯愛...
1087	想昨夜	黛玉	竟不	109 候芳魂五兒承錯愛...
1088	為著	黛玉	的事	109 候芳魂五兒承錯愛...
1089	他為	黛玉	之事	109 候芳魂五兒承錯愛...
1090	是為	黛玉	了。	109 候芳魂五兒承錯愛...
1091	便想起	黛玉	來，	115 惑偏私惜春矢素志...
1092	見是	黛玉	的形容	116 得通靈幻境悟仙緣...
1093	，想起	黛玉	一陣心酸	118 記微嫌勇兄欺弱女...
1094	自送	黛玉	的靈	120 甄士隱詳說太虛情...
1095	正室	夫人	了。	002 賈夫人仙逝揚州城...
1096	賈氏	夫人	一疾	002 賈夫人仙逝揚州城...
1097	，太	夫人	尚在	002 賈夫人仙逝揚州城...
1098	老爹的	夫人	王氏，	002 賈夫人仙逝揚州城...
1099	公之	夫人	，即	002 賈夫人仙逝揚州城...
1100	政老爹	夫人	王氏之內	002 賈夫人仙逝揚州城...
1101	他令	夫人	之後，	002 賈夫人仙逝揚州城...
1102	稱頌他	夫人	的，	002 賈夫人仙逝揚州城...
1103	」王	夫人	道：	003 金陵城起復賈雨村...
1104	」王	夫人	一笑	003 金陵城起復賈雨村...
1105	」邢	夫人	答應了	003 金陵城起復賈雨村...

匯出 搜尋到2539筆結果

上一頁 下一頁

5. 以 “” 括住表示緊鄰出現的詞組

CORPRO Alpha E:\Workspace\test\CORPRO Workspace\

設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性

搜尋詞: "我的" 執行 左數: 2 右數: 2 文章 ☐ 同類詞 ☐ 停用詞

#	左方出現詞	關鍵詞	右方出現詞	來源文章
1	租了	我的	房舍居住	004 薄命女偏逢薄命...
2	:	我的	小名這裏	005 開生面夢演紅樓...
3	:	我的	姥姥,	006 賈寶玉初試雲雨...
4	, 偏	我的	就是好	006 賈寶玉初試雲雨...
5	:	我的	娘啊	006 賈寶玉初試雲雨...
6	:	我的	嬖子,	006 賈寶玉初試雲雨...
7	不報	我的	恩,	007 送宮花周瑞嘆英...
8	:	我的	菩薩哥兒	008 薛寶釵小恙梨香...
9	說與	我的	小兒	008 薛寶釵小恙梨香...
10	天,	我的	兒,	008 薛寶釵小恙梨香...
11	怎麼瞧	我的	了呢	008 薛寶釵小恙梨香...
12	真與	我的	是一	008 薛寶釵小恙梨香...
13	取了	我的	斗篷來	008 薛寶釵小恙梨香...
14	怕,	我的	兒!	008 薛寶釵小恙梨香...
15	知道是	我的	, 偏	008 薛寶釵小恙梨香...
16	這就是	我的	意思,	009 戀風流情友入家...
17	講了	我的	門!	009 戀風流情友入家...
18	人欺侮	我的	。不如	009 戀風流情友入家...
19	。依	我的	主意,	009 戀風流情友入家...
20	都是	我的	嘴快	010 金寡婦貪利權受...

匯出 搜尋到406筆結果 上一頁 下一頁

七、停用詞

功能

1. 直接輸入新增

詞典 停用詞 同類詞 文檔檢視 詮釋資料欄位

的 新增 刪除

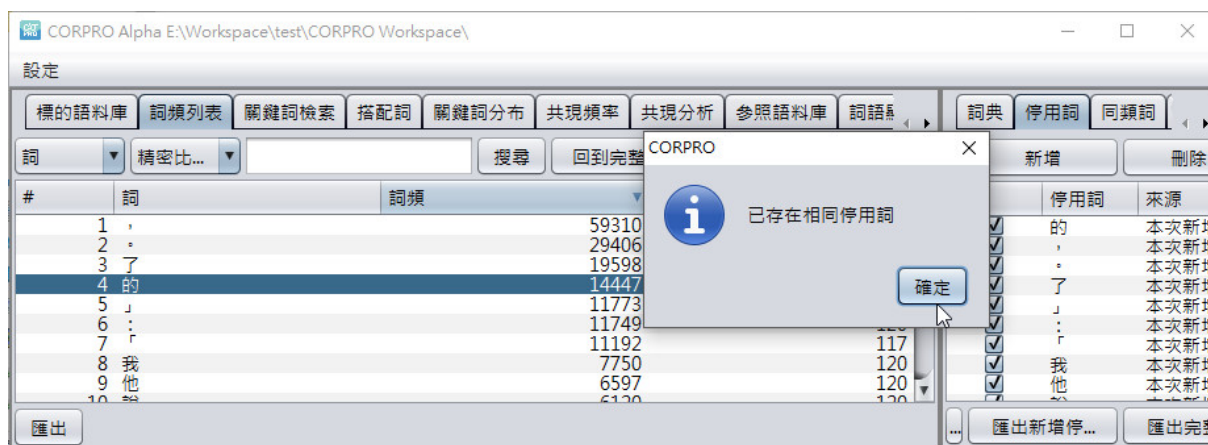
啟用	停用詞	來源
☒	的	本次新增

匯入 匯出新增停用詞 匯出完整列表

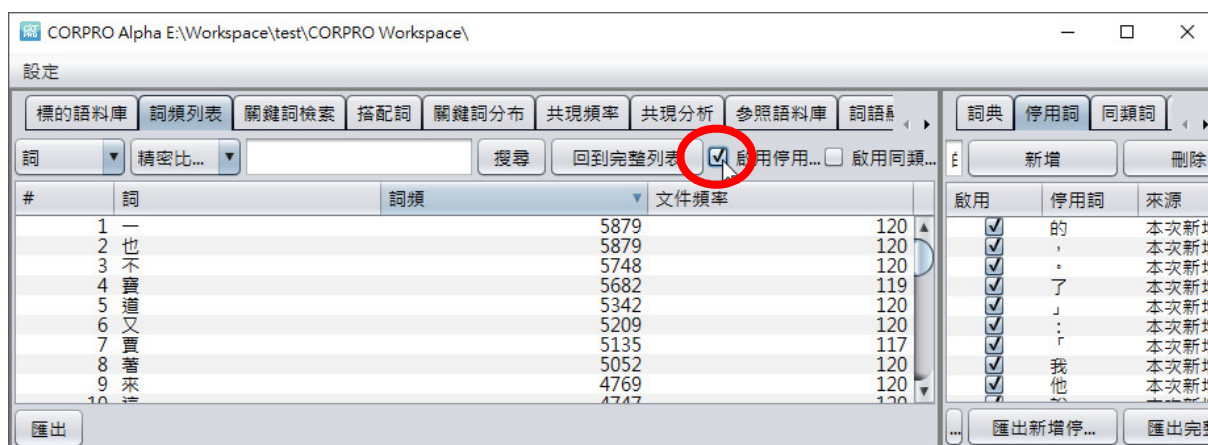
2. 可從詞頻列表一次拖曳多筆不連續詞新增至停用詞列表



3. 重複新增同一詞會被擋並出現警告訊息



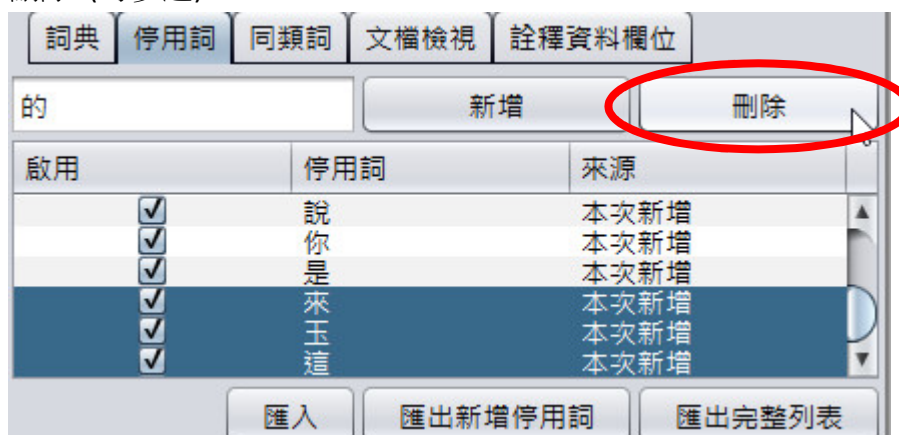
4. 詞頻列表→啟用停用詞 (停用詞不會出現)



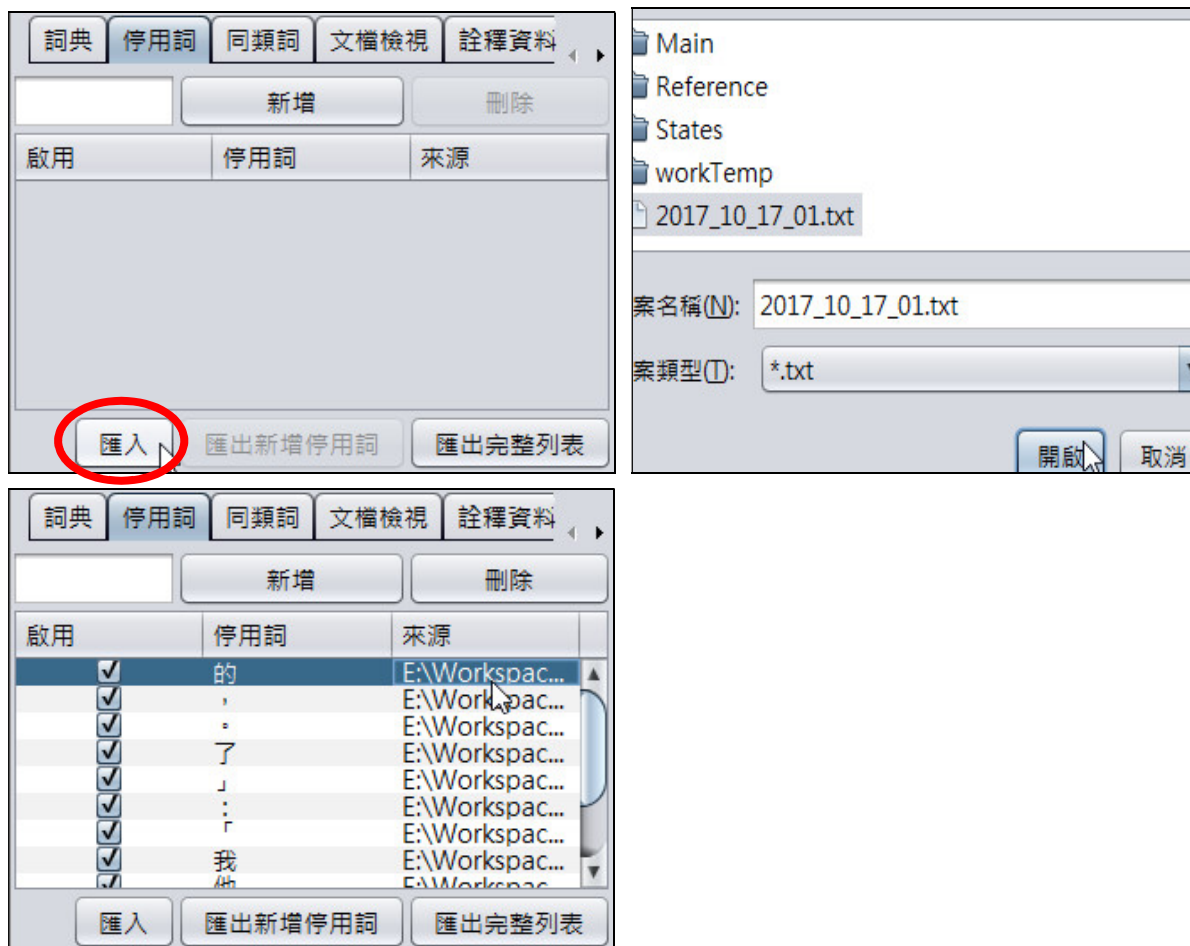
5. 停用詞的 checkbox 取消勾選可暫不視為停用詞 (以詞頻列表驗證)



6. 刪除 (可多選)



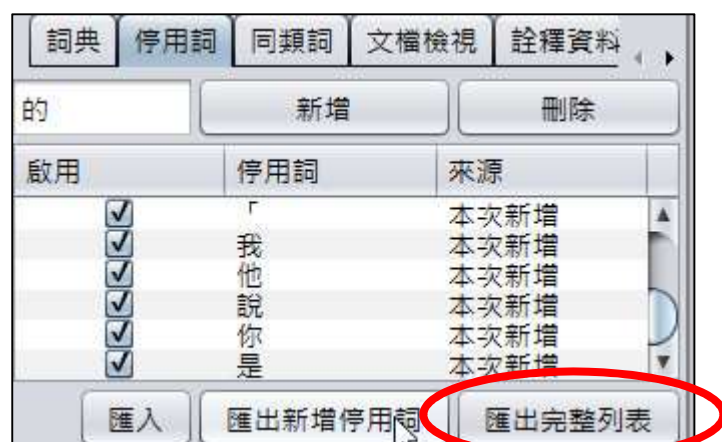
7. 匯入 (已存在之停用詞列表、會標出來源)



8. 匯出本次新增 (完成後來源會變)



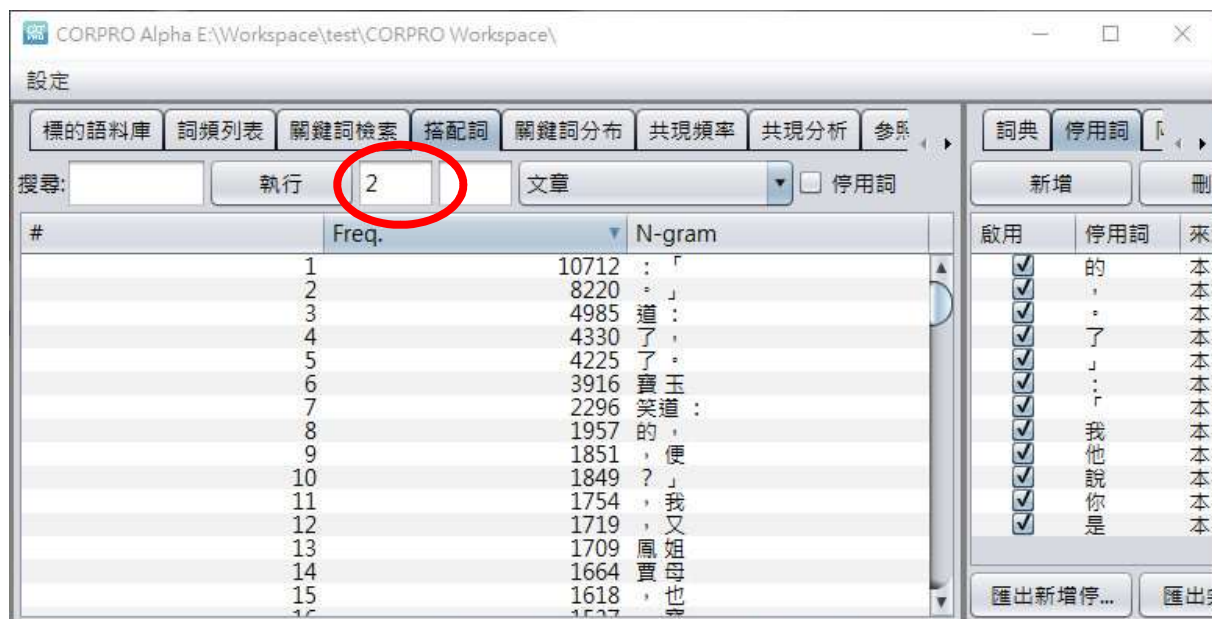
9. 匯出完整列表 (完成後來源會變)



八、搭配詞

功能

1. 僅輸入左詞數 n ，表示列出語料庫中所有的 n -grams



2. 輸入查詢詞 **w** 以及左詞數 **n**，表示列出所有包含有 **w** 的 **n**-grams

設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照

搜尋: 黛玉 執行 2 文章 ☐ 停用詞

#	Freq.	N-gram
1	358	「黛玉
2	200	。黛玉
3	141	黛玉道
4	98	，黛玉
5	85	黛玉笑道
6	56	黛玉，
7	47	黛玉，的
8	35	見黛玉
9	34	黛玉聽
10	33	黛玉便
11	32	著黛玉
12	26	黛玉又
13	24	了黛玉
14	24	黛玉。
15	23	黛玉也
16	22	黛玉

詞典 停用詞

新增 刪除

啟用 停用詞 來源

匯出新增停... 匯出...

3. 輸入查詢詞 **w**、左詞數 **n** 以及右詞數 **k**，表示列出 **w** 前 **n** 個詞、後 **k** 個詞的 **(n+k+1)**-grams

設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照

搜尋: 黛玉 執行 2 3 文章 ☐ 停用詞

#	Freq.	N-gram
1	83	。」「黛玉道：「
2	48	。」「黛玉笑道：「
3	28	？」「黛玉道：「
4	17	。」「黛玉聽了，
5	13	？」「黛玉笑道：「
6	5	！」「黛玉道：「
7	3	。」「黛玉冷笑道：「
8	3	。」「黛玉聽了這
9	3	。」「黛玉點點頭兒，
10	2	。」「黛玉笑道：「
11	2	。」「黛玉聽了，
12	2	。」「黛玉忙道：「
13	2	。」「黛玉聽說，便
14	2	。」「黛玉，湘雲
15	2	。」「黛玉聽了這
16	2	。」「黛玉

詞典 停用詞

新增 刪除

啟用 停用詞 來源

匯出新增停... 匯出...

4. 啟用停用詞的時候，包含任一停用詞的 **n**-gram 將不會被列出

設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照

搜尋: 黛玉 執行 2 3 文章 ☒ 停用詞

#	Freq.	N-gram
1	1	眾人見黛玉年貌雖
2	1	嬌嬌引黛玉進東房門
3	1	們讓黛玉炕上坐
4	1	忙攜黛玉從後房門
5	1	遂攜黛玉穿過一個東西
6	1	因問黛玉念何書
7	1	哥陪侍黛玉在碧紗
8	1	見裡面黛玉和鵝哥
9	1	日與黛玉迎春姊妹等一
10	1	玉和黛玉二人之下
11	1	故比黛玉大得下人
12	1	因與黛玉同隨賈
13	1	寶玉黛玉二人這
14	1	便到黛玉房中去
15	1	知此時黛玉不在
16	1	黛玉

詞典 停用詞

新增 刪除

啟用 停用詞 來源

匯出新增停... 匯出...

5. 一樣有四種邊界條件 (只比較 “文章” 和 “標點”)

CORPRO Alpha E:\Workspace\test\CORPRO Workspace\

設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照

搜尋: 黛玉 執行 2 3 文章 停用詞

#	Freq.	N-gram
1	2	釵, 黛玉, 湘雲
2	2	釵, 黛玉, 寶玉
3	2	琴, 黛玉, 湘雲
4	1	乳名黛玉, 年方
5	1	女學生黛玉, 身體方
6	1	依附黛玉, 而行日
7	1	且說黛玉下榻。林黛玉
8	1	扶, 黛玉也哭個
9	1	涕泣, 黛玉忙起身迎上來
10	1	飾。黛玉忙起身迎上來
11	1	眾人見黛玉連忙起身接見
12	1	開。黛玉忙陪笑見禮
13	1	鳳。黛玉忙陪笑見禮
14	1	忙攜黛玉之手, 房屋
15	1	中。黛玉度其房屋

啟用 停用詞 來源

新增 刪除

匯出新增停... 匯出...

CORPRO Alpha E:\Workspace\test\CORPRO Workspace\

設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照

搜尋: 黛玉 執行 2 3 標點 停用詞

#	Freq.	N-gram
1	1	眾人見黛玉年貌雖
2	1	嫵嫵引黛玉進東房門
3	1	忙攜黛玉從後房門
4	1	遂攜黛玉穿過一個東西
5	1	哥陪侍黛玉在碧紗
6	1	見裡面黛玉和鸚哥
7	1	日與黛玉迎春姊妹等一
8	1	玉和黛玉二人之下
9	1	故比黛玉大得下人
10	1	因與黛玉同隨賈
11	1	寶玉黛玉二人這邊
12	1	便到黛玉房中去
13	1	知此時黛玉不在自己
14	1	走來與黛玉送小手
15	1	忙至黛玉房中來

啟用 停用詞 來源

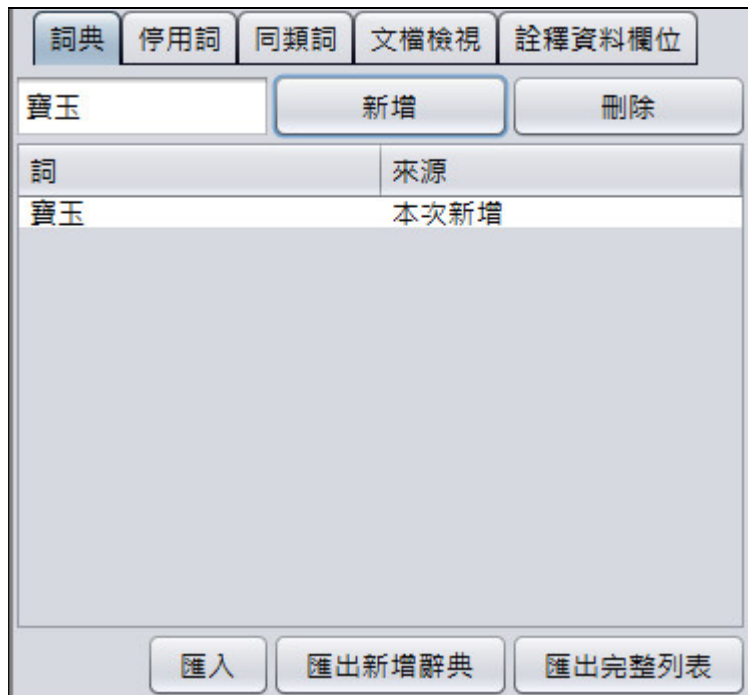
新增 刪除

匯出新增停... 匯出...

九、詞典

功能

1. 直接輸入新增



詞	來源
寶玉	本次新增

2. 可從搭配詞一次拖曳多筆不連續詞新增至停用詞列表 (請舉有意義的組合為例)

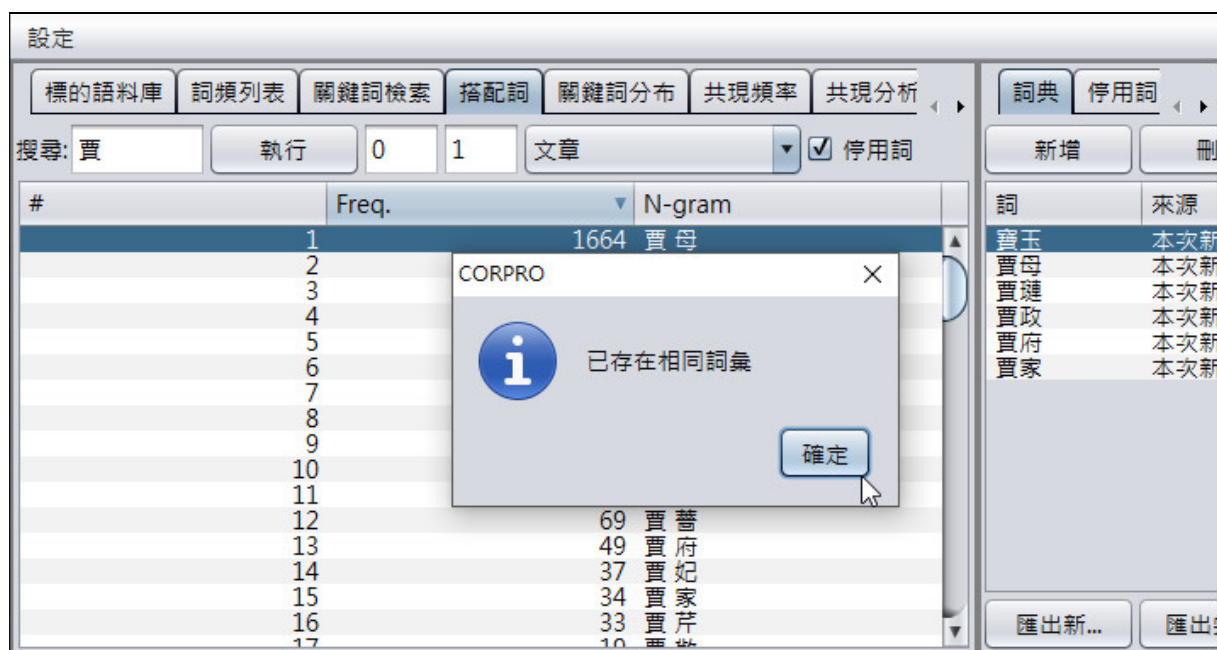


#	Freq.	N-gram
1	1664	賈母
2	766	賈璉
3	745	賈政
4	387	賈珍
5	230	賈蓉
6	194	賈赦
7	191	賈芸
8	170	賈政道
9	151	賈環
10	93	賈蘭
11	72	賈瑞
12	69	賈薈
13	49	賈府
14	37	賈妃
15	34	賈家
16	33	賈芹
17	10	賈

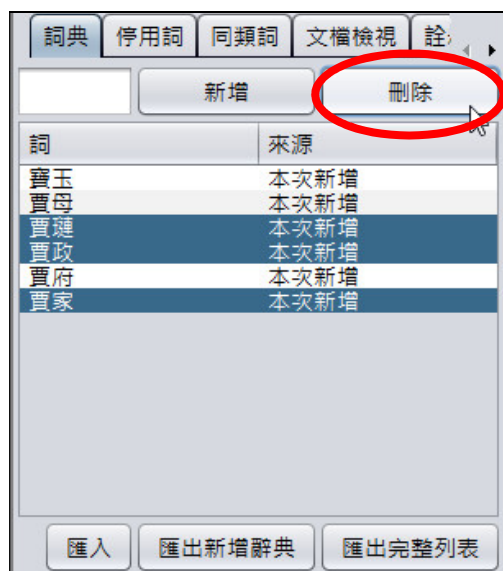


詞	來源
寶玉	本次新增
賈母	本次新增
賈璉	本次新增
賈政	本次新增
賈珍	本次新增
賈蓉	本次新增
賈府	本次新增
賈家	本次新增

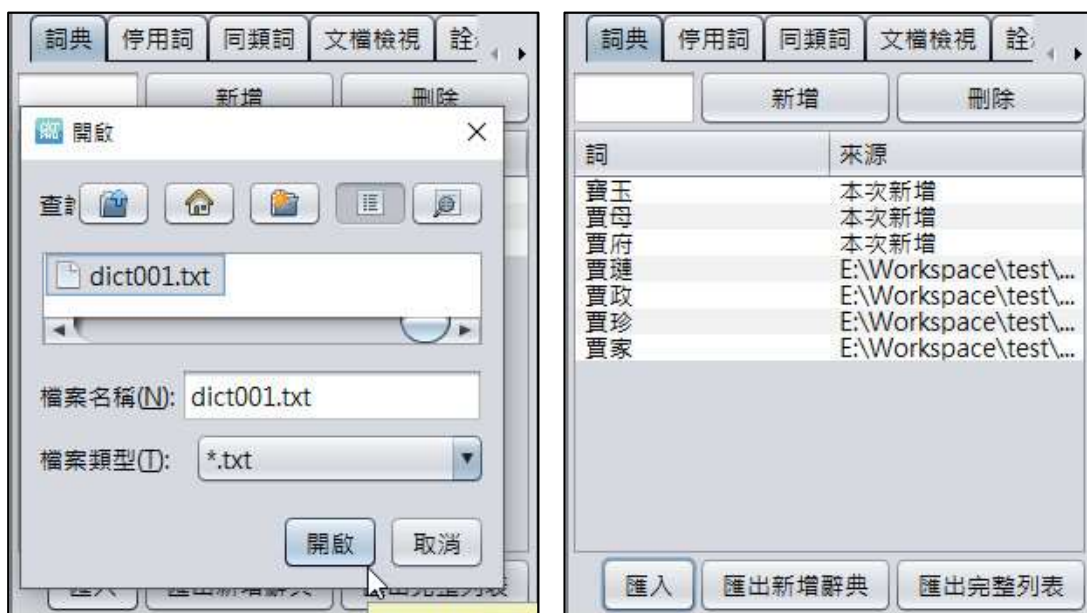
3. 重複新增同一詞會被擋並出現警告訊息



4. 刪除 (可多選)



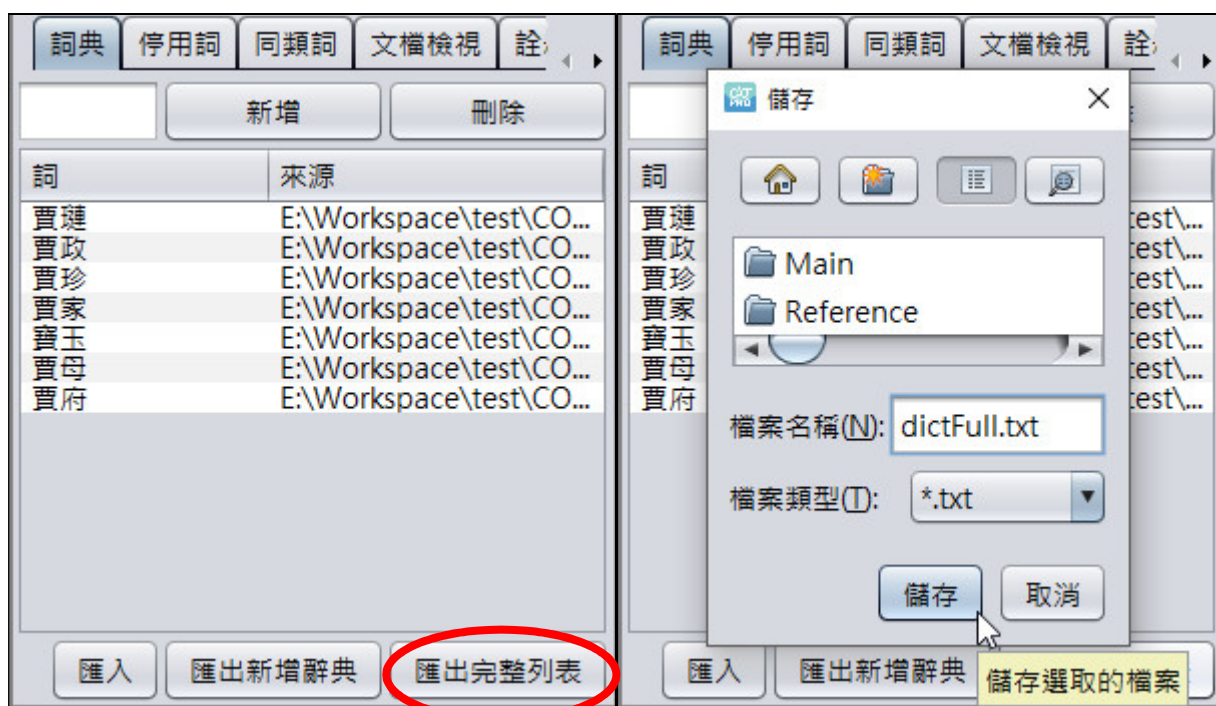
5. 匯入 (已存在之停用詞列表、會標出來源)



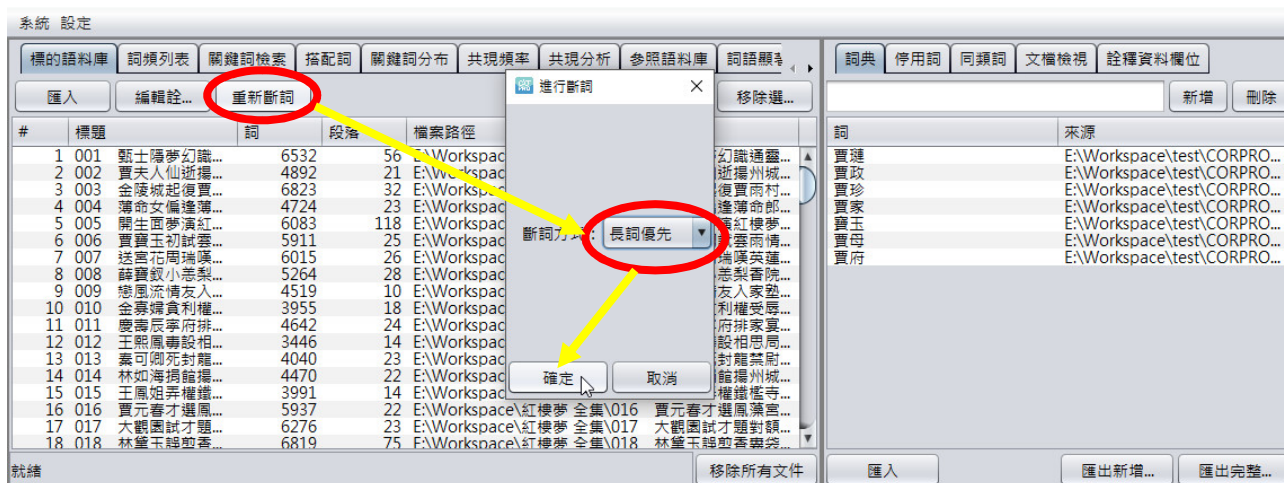
6. 匯出本次新增 (完成後來源會變)



7. 匯出完整列表 (完成後來源會變)



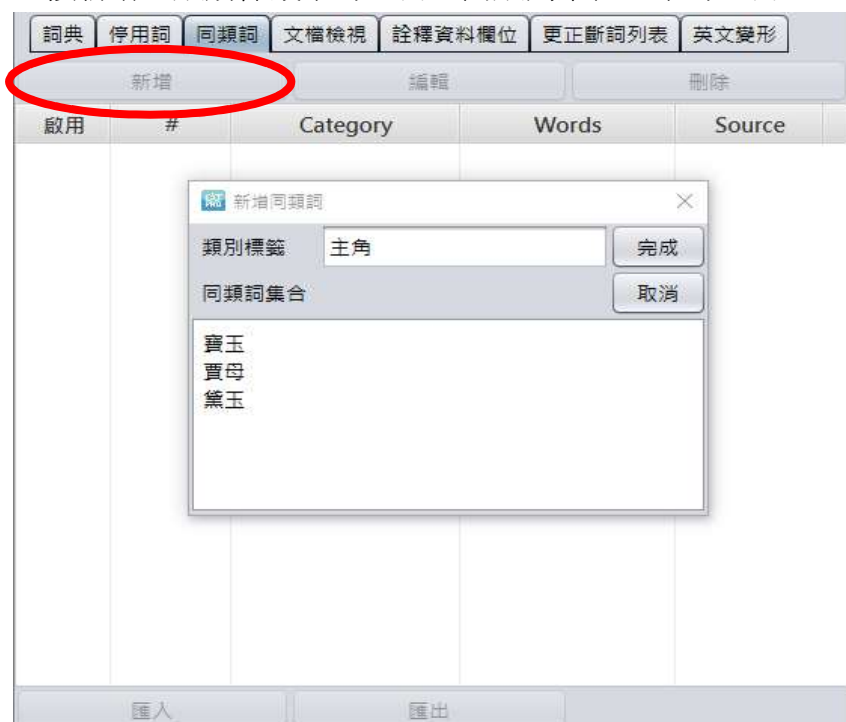
8. 標的語料庫→重新斷詞 (詞典中的新詞彙加入了斷詞系統，同時會重斷標的語料庫及參考語料庫，可用詞頻列表印證新詞的出現)



十、同類詞

功能

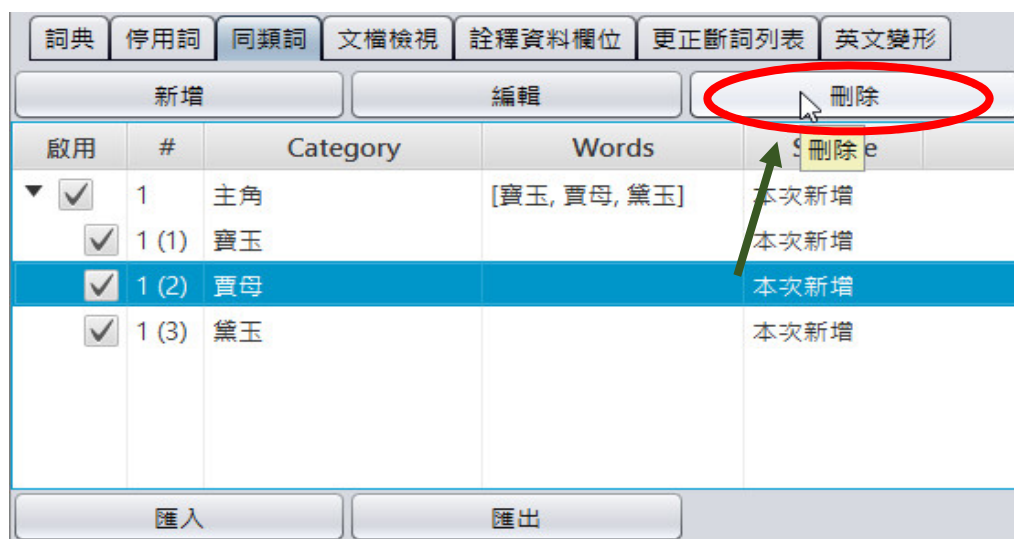
1. 直接新增，類別標籤不可空白，同類詞集合也不可空白



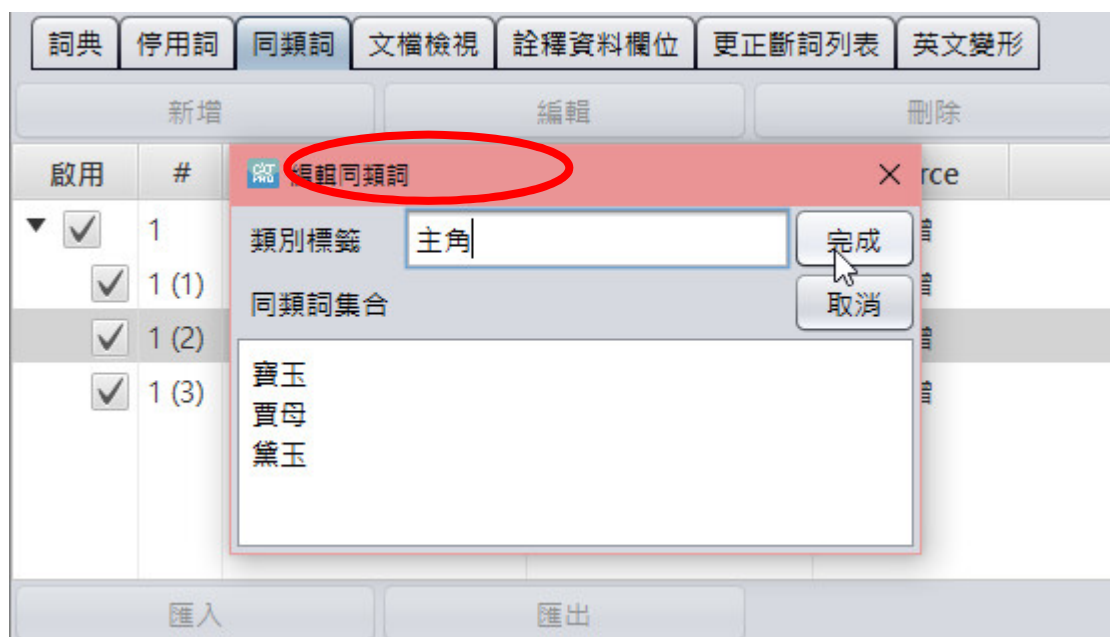
2. 可刪除整個類別



3. 也可刪除類別中某一個詞



4. 如果想新增刪除某個類別裡的詞，可按編輯



5. 可展開各同義詞集合

詞典	停用詞	同類詞	文檔檢視	詮釋資料欄位	更正斷詞列表	英文變形
新增			編輯		刪除	
啟用	#	Category	Words	Source		
☒	1	主角	[寶玉, 賈母, 黛玉]	本次新增		
☒	1 (1)	寶玉		本次新增		
☒	1 (2)	賈母		本次新增		
☒	1 (3)	黛玉		本次新增		
匯入			匯出			

十一、詞頻列表與同類詞功能

1. 若在詞頻列表啟用同類詞，則在同一個集合裡的同類詞的詞頻會加總，文件頻率也會合併

系統 設定

標的語料庫

詞頻列表

關鍵詞檢索

搭配詞

關鍵詞分布

共現頻率

共現分析

參照語料庫

詞

精密...

搜尋

回到完整列...

啟用停...

啟用同...

#	詞	詞頻	文件頻率	同類詞
43	再	1034	120	
44	些	965	120	
45	下	937	120	
46	說道	916	120	
47	起來	646	120	
48	▶ 主角	2775	119	賈母[1681,100],黛玉[1...
49	兒	4284	119	
50	▶ 稱謂	4230	119	丫頭[992,112],老太太[...
51	之	1770	119	
52	回	1269	119	
53	我們	1200	119	

匯出

詞典

停用詞

同類詞

文檔檢視

詮釋資料權位

新增

編輯

刪除

#	啟用	Category	Words	Source
1	☒	▶ 主角	[寶玉, 黛玉, ...	本次新增
2	☒	▶ 稱謂	[老太太, 太...	本次新增

匯入

匯出

2. 可選用或取消整個同義詞集合 (目前需取消再重新啟用才能看到變化)

系統 設定

標的語料庫

詞頻列表

關鍵詞檢索

搭配詞

關鍵詞分布

共現頻率

共現分析

參照語料庫

詞

精密...

搜尋

回到完整列...

☒ 啟用停...

☒ 啟用同...

#	詞	詞頻	文件頻率	同類詞
46	說道	916	120	
47	起來	646	120	
48	主角	6719	119	寶玉[3944.117], 賈母[1...
49	兒	4284	119	
50	之	1770	119	
51	回	1269	119	
52	我們	1200	119	
53	話	1098	119	
54	把	1029	119	
55	得	1021	119	
56	如今	1012	119	

匯出

詞典

停用詞

同類詞

文檔檢視

詮釋資料欄位

新增

編輯

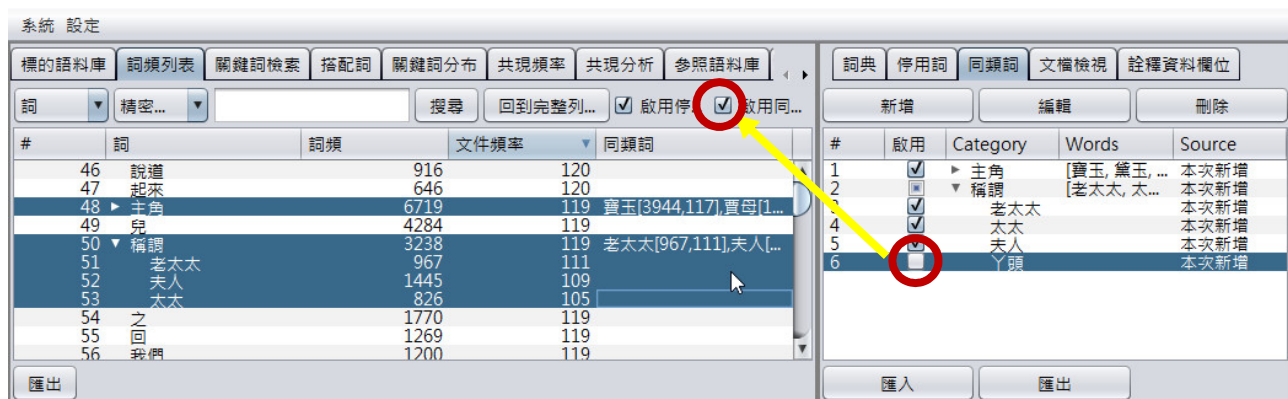
刪除

#	啟用	Category	Words	Source
1	☒	▶ 主角	[寶玉, 黛玉, ...	本次新增
2	☒	▶ 稱謂	[老太太, 太...	本次新增

匯入

匯出

3. 可選用或取消個別詞（目前需取消再重新啟用才能看到變化）

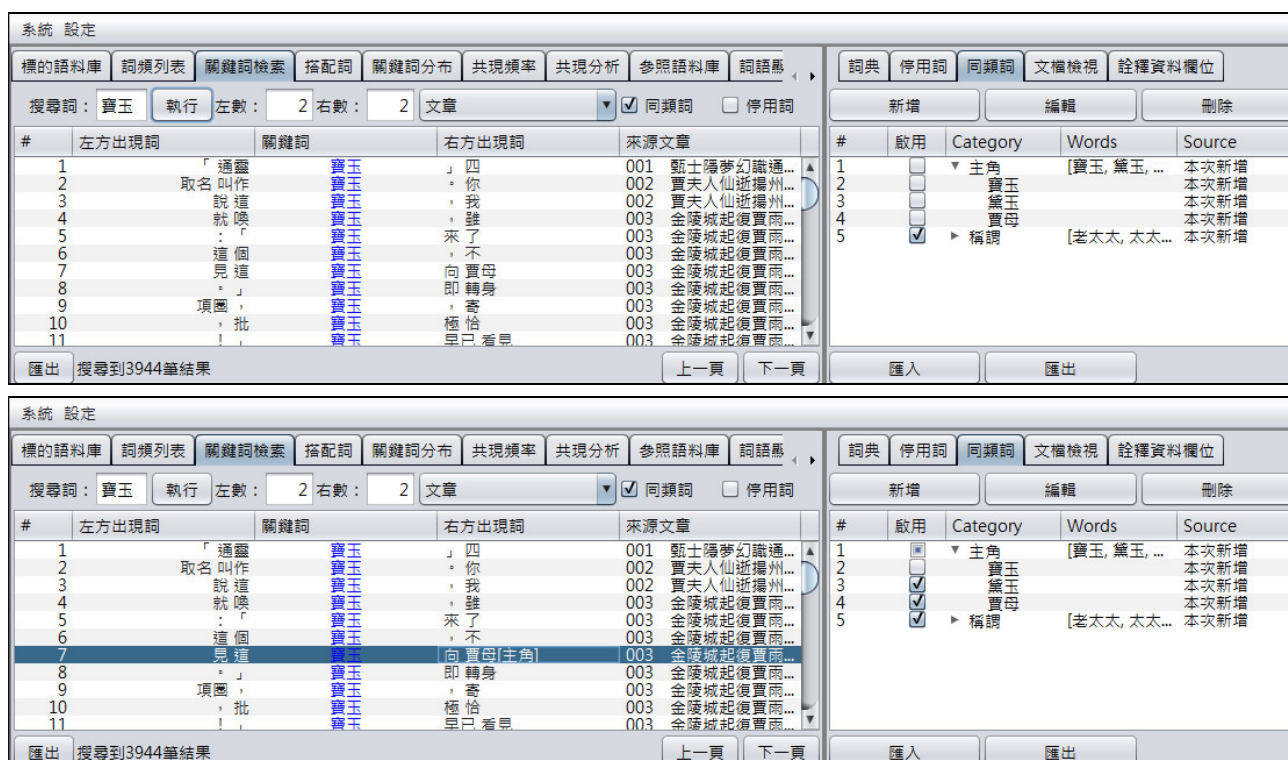


十二、 關鍵詞查詢與同類詞功能

1. 若在詞頻列表啟用同類詞，出現在同類詞集合裡的詞頻會加註它的類別標籤



2. 可選用或取消整個同義詞集合、或選用取消個別詞



- 若查詢詞本身就是類別標籤，則會查出該類別裡的所有詞

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語屬性

搜尋詞：主角 執行 左數：2 右數：2 文章 ☒ 同類詞 ☐ 停用詞

#	左方出現詞	關鍵詞	右方出現詞	來源文章
108	自此	寶玉(主角)	視 襲人	006 寶寶玉初試雲雨...
109	襲人侍	寶玉(主角)	更為 盡心	006 寶寶玉初試雲雨...
110	只留	寶玉(主角)	黛玉(主角) 二	007 送宮花周瑞漢英...
111	卻在	寶玉(主角)	房中大	007 送宮花周瑞漢英...
112	「就	寶玉(主角)	聽說，	007 送宮花周瑞漢英...
113	只就	寶玉(主角)	手 中看	007 送宮花周瑞漢英...
114	言語，	寶玉(主角)	便問道	007 送宮花周瑞漢英...
115	「	寶玉(主角)	道：	007 送宮花周瑞漢英...
116	「	寶玉(主角)	聽了	007 送宮花周瑞漢英...
117	賈母(主角)	寶玉(主角)	聽了	007 送宮花周瑞漢英...
118	攔了	寶玉(主角)	同入	007 送宮花周瑞漢英...

匯入 搜尋到6719筆結果 上一頁 下一頁

詞典 停用詞 同類詞 文檔檢視 詮釋資料欄位

新增 編輯 刪除

#	啟用	Category	Words	Source
1	☒	主角	[寶玉, 黛玉, ...	本次新增
2	☒	寶玉		本次新增
3	☒	黛玉		本次新增
4	☒	賈母		本次新增
5	☒	稱謂	[老太太, 太太...	本次新增

匯入 匯出

十三、 詮釋資料欄位

功能

- 在右邊「編輯詮釋資料欄位」頁籤可新增詮釋資料欄位 (註：新增後無法刪除)。(範例改用總統就職演說，新增「總統」、「產生方式」兩種欄位)

詞典 停用詞 同類詞 文檔檢視 詮釋資料欄位

產生方式

啟用 Word

☐ 總統

新增

- 新增後在「標的語料庫」表格中會出現相對應的欄位

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語屬性 字串類

匯入 編輯詮釋... 重新斷詞 移除選擇...

#	標題	詞	段落	總統	產生方式	檔案路徑
1	1948蔣中正.docx	2452	9			E:\Workspace\總統就職演說 194...
2	1954蔣中正.docx	1702	14			E:\Workspace\總統就職演說 194...
3	1960蔣中正.docx	1872	11			E:\Workspace\總統就職演說 194...
4	1966蔣中正.docx	2557	16			E:\Workspace\總統就職演說 194...
5	1972 蔣中正.docx	692	7			E:\Workspace\總統就職演說 194...
6	1978 蔣經國.docx	1330	17			E:\Workspace\總統就職演說 194...
7	1984 蔣經國.docx	735	12			E:\Workspace\總統就職演說 194...
8	1990 李登輝.docx	1871	14			E:\Workspace\總統就職演說 194...
9	1996 李登輝.docx	2674	38			E:\Workspace\總統就職演說 194...
10	2000 陳水扁.docx	3315	54			E:\Workspace\總統就職演說 194...
11	2004 陳水扁.docx	3460	46			E:\Workspace\總統就職演說 194...
12	2008 馬英九.docx	2325	29			E:\Workspace\總統就職演說 194...
13	2012 馬英九.docx	2325	29			E:\Workspace\總統就職演說 194...

就緒 移除所有文件

詞典 停用詞 同類詞 文檔檢視 詮釋資料欄位

新增

啟用 Word

☒ 總統

☒ 產生方式

- 可選用或取消詮釋資料欄位，標的語料庫表格的相對應欄位會出現或消失

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語屬性 字串類

匯入 編輯詮釋... 重新斷詞 移除選擇...

#	標題	詞	段落	總統	檔案路徑
1	1948蔣中正.docx	2452	9		E:\Workspace\總統就職演說 1948 2016\1948蔣...
2	1954蔣中正.docx	1702	14		E:\Workspace\總統就職演說 1948 2016\1954蔣...
3	1960蔣中正.docx	1872	11		E:\Workspace\總統就職演說 1948 2016\1960蔣...
4	1966蔣中正.docx	2557	16		E:\Workspace\總統就職演說 1948 2016\1966蔣...
5	1972 蔣中正.docx	692	7		E:\Workspace\總統就職演說 1948 2016\1972 蔣...
6	1978 蔣經國.docx	1330	17		E:\Workspace\總統就職演說 1948 2016\1978 蔣...
7	1984 蔣經國.docx	735	12		E:\Workspace\總統就職演說 1948 2016\1984 蔣...
8	1990 李登輝.docx	1871	14		E:\Workspace\總統就職演說 1948 2016\1990 李...
9	1996 李登輝.docx	2674	38		E:\Workspace\總統就職演說 1948 2016\1996 李...
10	2000 陳水扁.docx	3315	54		E:\Workspace\總統就職演說 1948 2016\2000 陳...
11	2004 陳水扁.docx	3460	46		E:\Workspace\總統就職演說 1948 2016\2004 陳...
12	2008 馬英九.docx	2325	29		E:\Workspace\總統就職演說 1948 2016\2008 馬...
13	2012 馬英九.docx	2325	29		E:\Workspace\總統就職演說 1948 2016\2012 馬...

就緒 移除所有文件

詞典 停用詞 同類詞 文檔檢視 詮釋資料欄位

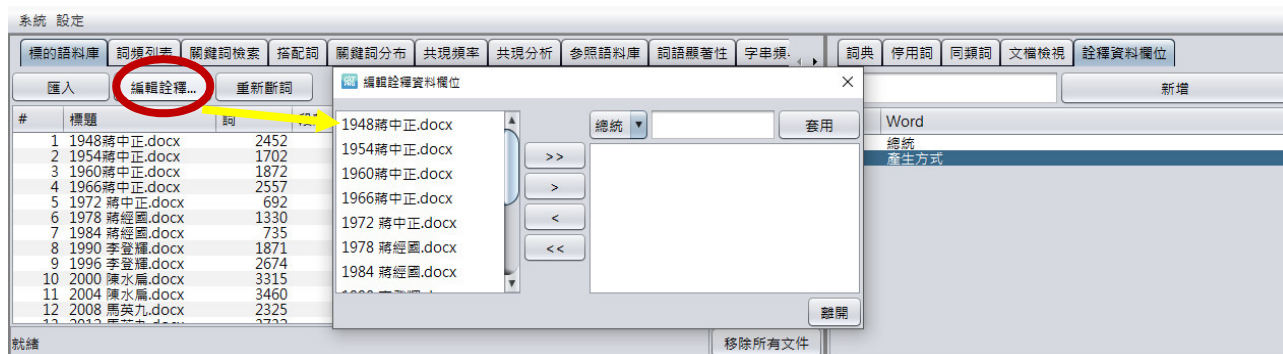
新增

啟用 Word

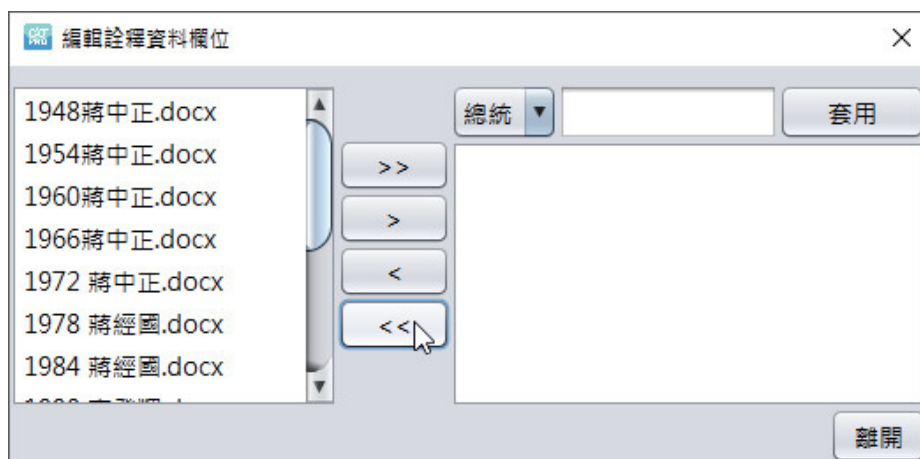
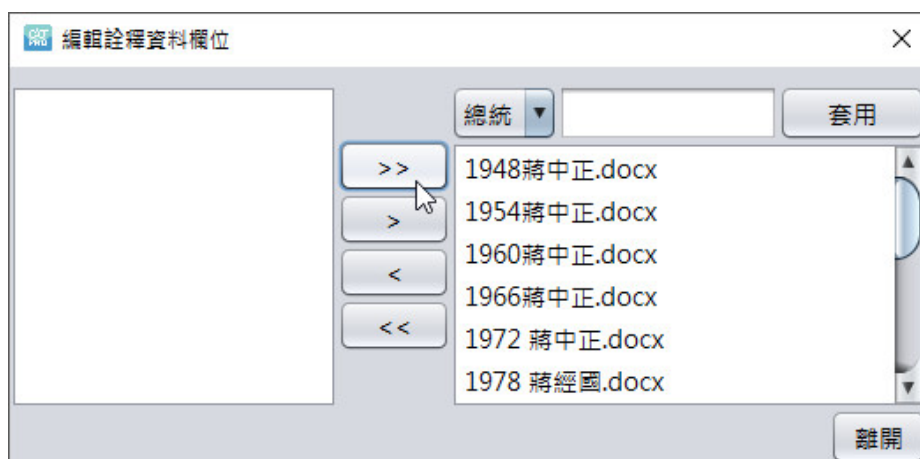
☒ 總統

☐ 產生方式

4. 在「標的語料庫」按「編輯詮釋資料」可大批設定詮釋資料



5. 按 >> 可加進全部檔案；按 << 可移出全部檔案；
按 > 可加進選取的檔案；按 < 可移出選取的檔案



6. 選擇要套用的欄位，輸入這群檔案的詮釋資料值，按「套用」來完成。(已給定詮釋資料的檔案會變成綠色字並放在左方列表最後)



7. 回到「標的語料庫」表格可以看到那些檔案的對應欄位已有值

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性 字串頻

匯入 編輯詮釋... 重新斷詞 移除選擇...

#	標題	詞	段落	總統	檔案路徑
1	1948蔣中正.docx	2452	9	蔣中正	E:\Workspace\總統就職演說 1948 2016\1948蔣...
2	1954蔣中正.docx	1702	14	蔣中正	E:\Workspace\總統就職演說 1948 2016\1954蔣...
3	1960蔣中正.docx	1872	11	蔣中正	E:\Workspace\總統就職演說 1948 2016\1960蔣...
4	1966蔣中正.docx	2557	16	蔣中正	E:\Workspace\總統就職演說 1948 2016\1966蔣...
5	1972 蔣中正.docx	692	7	蔣中正	E:\Workspace\總統就職演說 1948 2016\1972 蔣...
6	1978 蔣經國.docx	1330	17		E:\Workspace\總統就職演說 1948 2016\1978 蔣...
7	1984 蔣經國.docx	735	12		E:\Workspace\總統就職演說 1948 2016\1984 蔣...
8	1990 李登輝.docx	1871	14		E:\Workspace\總統就職演說 1948 2016\1990 李...
9	1996 李登輝.docx	2674	38		E:\Workspace\總統就職演說 1948 2016\1996 李...
10	2000 陳水扁.docx	3315	54		E:\Workspace\總統就職演說 1948 2016\2000 陳...
11	2004 陳水扁.docx	3460	46		E:\Workspace\總統就職演說 1948 2016\2004 陳...
12	2008 馬英九.docx	2325	29		E:\Workspace\總統就職演說 1948 2016\2008 馬...
13	2012 馬英九.docx	2325	46		E:\Workspace\總統就職演說 1948 2016\2012 馬...

就緒 移除所有文件

8. 在標的語料庫表格的詮釋資料欄點兩下，可以直接編輯詮釋資料值

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性 字串頻

匯入 編輯詮釋... 重新斷詞 移除選擇...

#	標題	詞	段落	總統	檔案路徑
1	1948蔣中正.docx	2452	9	蔣中正	E:\Workspace\總統就職演說 1948 2016\1948蔣...
2	1954蔣中正.docx	1702	14	蔣中正	E:\Workspace\總統就職演說 1948 2016\1954蔣...
3	1960蔣中正.docx	1872	11	蔣中正	E:\Workspace\總統就職演說 1948 2016\1960蔣...
4	1966蔣中正.docx	2557	16	蔣中正	E:\Workspace\總統就職演說 1948 2016\1966蔣...
5	1972 蔣中正.docx	692	7	蔣中正	E:\Workspace\總統就職演說 1948 2016\1972 蔣...
6	1978 蔣經國.docx	1330	17		E:\Workspace\總統就職演說 1948 2016\1978 蔣...
7	1984 蔣經國.docx	735	12		E:\Workspace\總統就職演說 1948 2016\1984 蔣...
8	1990 李登輝.docx	1871	14		E:\Workspace\總統就職演說 1948 2016\1990 李...
9	1996 李登輝.docx	2674	38		E:\Workspace\總統就職演說 1948 2016\1996 李...
10	2000 陳水扁.docx	3315	54		E:\Workspace\總統就職演說 1948 2016\2000 陳...
11	2004 陳水扁.docx	3460	46		E:\Workspace\總統就職演說 1948 2016\2004 陳...
12	2008 馬英九.docx	2325	29		E:\Workspace\總統就職演說 1948 2016\2008 馬...
13	2012 馬英九.docx	2325	46		E:\Workspace\總統就職演說 1948 2016\2012 馬...

就緒 移除所有文件

十四、詞語趨勢

功能

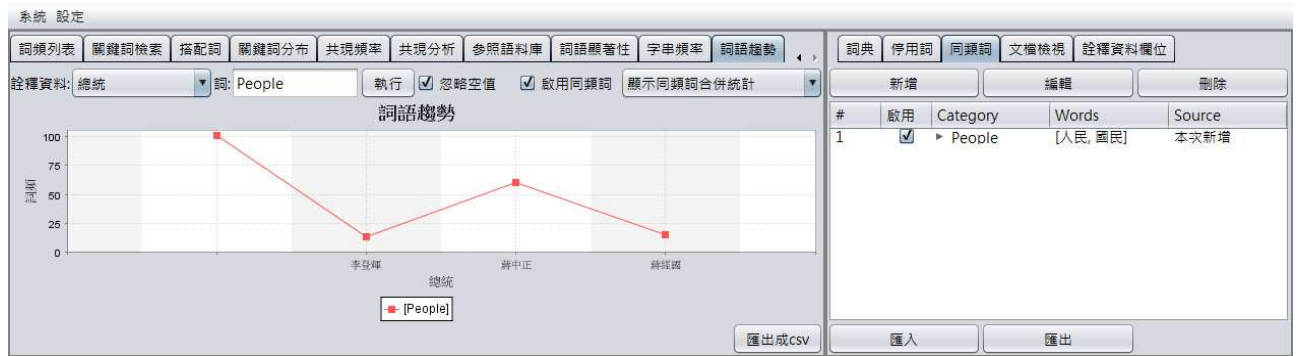
1. 先選擇要分析的詮釋資料欄位，再輸入查詢並按「執行」



2. 沒有設定詮釋資料 (亦即是空值) 的檔案會自成一類。若要忽略這群檔案可啟用空值



3. 若啟用同類詞、且查詢詞為類別標籤、在右上角選單選擇「顯示同義詞合併統計」時，集合中所有詞會合併顯示統計表



4. 在右上角選單選擇「顯示同義詞分別統計」，則集合中各個詞各自有統計線



5. 在右上角選單選擇「兩者皆顯示」，除了集合中各詞各自統計線外，也有合併的統計線

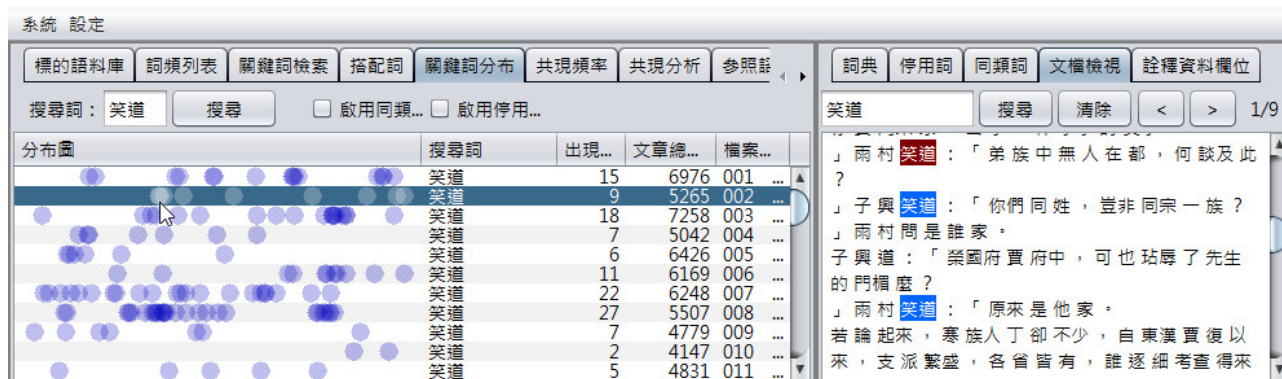


功能

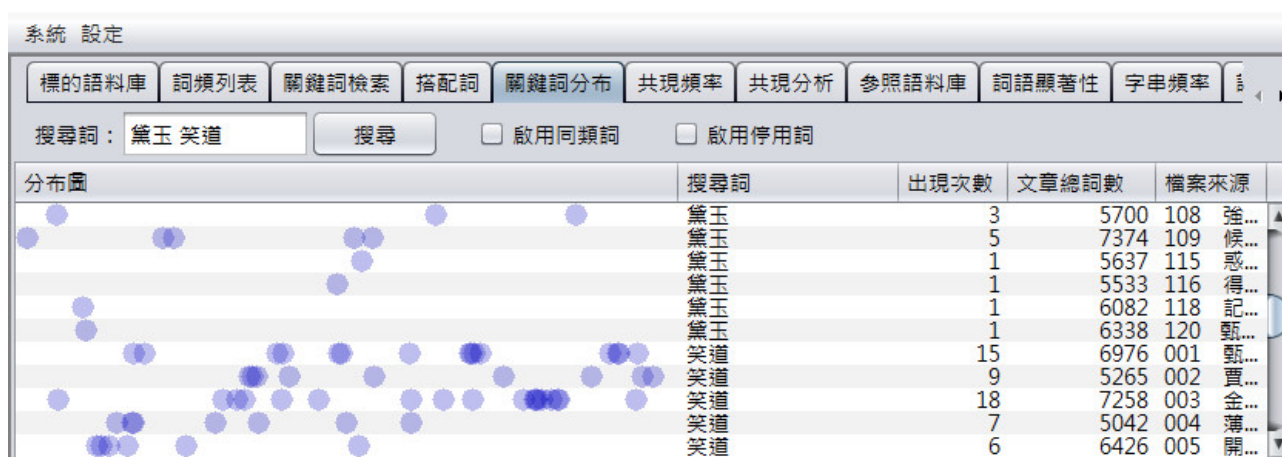
1. 輸入查詢並按「執行」



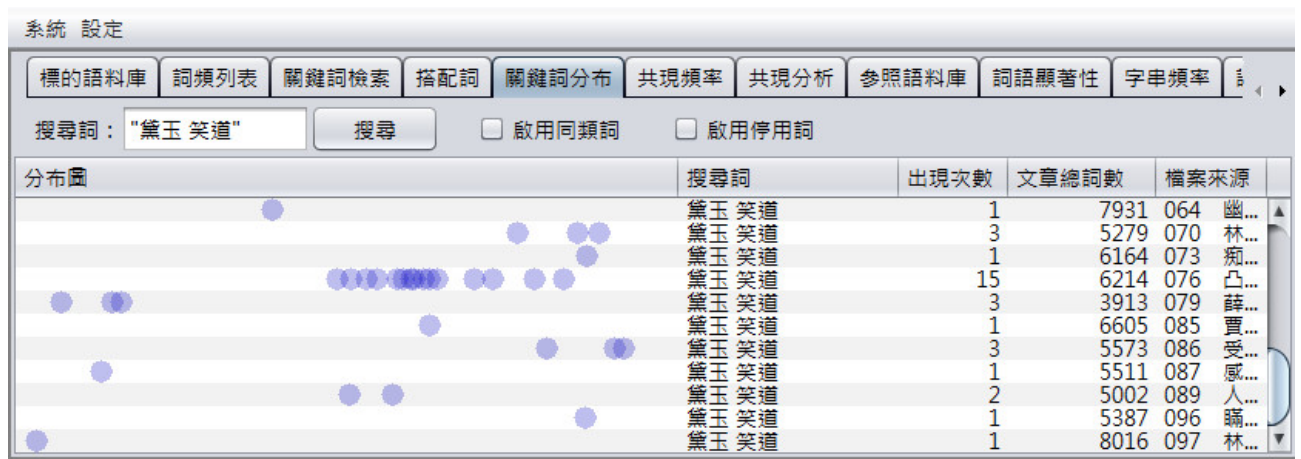
2. 在分佈圖條點兩下會跳至文檔檢視+查詢



3. 查詢兩個以上的詞、用空白隔開表示 OR (截圖時可把右邊切掉)



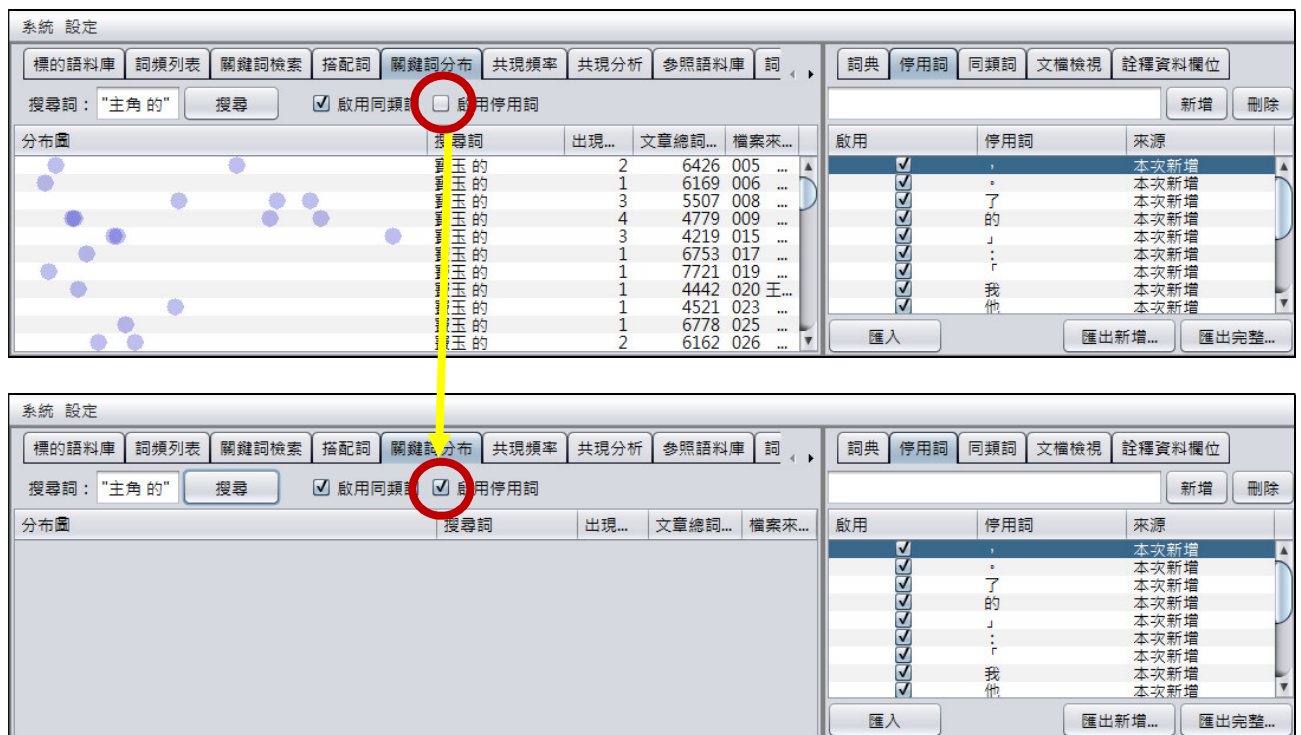
4. 以 "" 括住表示緊鄰出現的詞組



5. 啟用同類詞



6. 啟用停用詞 (停用詞不會出現)



十六、 共現頻率

功能

1. 輸入詞數並按「執行」(以 3+標點為例)

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著

搜尋詞： 執行 詞數： 3 標點 距離限制： 門檻值： ☐ 啟用停用詞 ☐ 啟用同類詞

#	Freq.	Co-Words
1	3072	」道：
2	1341	」了，
3	1320	」笑道：
4	1285	是的，
5	1232	了的，
6	1047	」了，
7	1044	」了的，
8	1031	」的，
9	1021	」我，
10	984	」說，
11	978	一了，

2. 可設定距離限制

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著

搜尋詞： 執行 詞數： 3 標點 距離限制： 5 門檻值： ☐ 啟用停用詞 ☐ 啟用同類詞

#	Freq.	Co-Words
1	2735	」道：
2	1242	」笑道：
3	878	了聽，
4	777	一了，
5	772	」著說
6	771	」說，
7	754	是的，
8	734	了的，
9	712	著說，
10	706	」了聽
11	653	」了，

3. 可設定門值

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著

搜尋詞： 執行 詞數： 3 標點 距離限制： 5 門檻值： 700 ☐ 啟用停用詞 ☐ 啟用同類詞

#	Freq.	Co-Words
1	2735	」道：
2	1242	」笑道：
3	878	了聽，
4	777	一了，
5	772	」著說
6	771	」說，
7	754	是的，
8	734	了的，
9	712	著說，
10	706	」了聽

4. 若有搜尋詞，只會出現含有搜尋詞的共現組合

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著

搜尋詞： 寶玉 執行 詞數： 3 標點 距離限制： 5 門檻值： ☐ 啟用停用詞 ☐ 啟用同類詞

#	Fr...	Co-Words
1	643	」寶玉：
2	606	寶玉道：
3	451	了寶玉，
4	432	」寶玉道
5	313	」寶玉，
6	274	寶玉笑道：
7	229	」了寶玉
8	197	」寶玉笑道
9	189	了寶玉聽
10	181	寶玉聽，
11	173	寶玉見，

5. 啟用停用詞時，共現組合中有停用詞的不會呈現

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性

搜尋詞： 寶玉 執行 詞數： 3 標點 距離限制： 5 門檻值： ☒ 啟用停用詞 ☐ 啟用同類詞

#	Freq.	Co-Words
1	40	只 寶玉 見
2	36	夫人 寶玉 王
3	29	寶玉 忙 道
4	27	又 寶玉 道
5	26	上 在 寶玉
6	25	向 寶玉 道
7	23	寶玉 聽 話
8	22	來 寶玉 見
9	22	寶玉 聽 這
10	22	寶玉 話 這
11	21	寶玉 笑 著

6. 多詞查詢

第一個範例以「寶玉 寶釵」做為查詢，代表想找兩個詞中至少有一個詞出現的組合。

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性 字串頻率

搜尋詞： 寶玉 寶釵 執行 詞數： 3 標點 距離限制： 門檻值： ☐ 不啟用同類詞

#	Freq.	Co-Words
1	882	寶玉 道：
2	681	」 寶玉：
3	634	」 寶玉 道
4	590	了 寶玉，
5	542	」 寶玉，
6	298	寶玉 笑：
7	282	寶玉 笑 道
8	279	寶釵 道：
9	254	」 了 寶玉
10	238	寶玉 的，
11	237	寶玉 見，
12	228	寶玉 聽，
13	221	」 寶玉 笑
14	201	」 寶釵：
15	194	了 寶玉 聽
16	190	」 寶釵 道
17	181	」 寶玉 聽
18	139	。 了 寶玉
19	130	寶釵 笑：

匯出 選取匯出 就緒 - 總資料數：2666007 上一頁 [1] ~ [200]

第二個範例以“寶玉 寶釵”做為查詢，因為查詢以雙括號夾註，代表兩個詞都要出現，因此會列出這兩個詞都有出現的組合。

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性 字串頻率 詞語

搜尋詞： "寶玉 寶釵" 執行 詞數： 3 標點 距離限制： 門檻值： 不啟用同類詞

#	Freq.	Co-Words
1	42	寶玉 寶釵，
2	10	了 寶玉 寶釵
3	9	寶玉 寶釵 見
4	8	寶玉 寶釵 的
5	8	。 寶玉 寶釵
6	6	」 寶玉 寶釵
7	4	寶玉 寶釵 說
8	4	在 寶玉 寶釵
9	3	只見 寶玉 寶釵
10	3	他 寶玉 寶釵
11	3	寶玉 寶釵 看
12	3	且說 寶玉 寶釵
13	3	因 寶玉 寶釵
14	3	寶玉 寶釵 心
15	3	又 寶玉 寶釵
16	3	來 寶玉 寶釵
17	3	寶玉 寶釵 都
18	3	寶玉 寶釵 雖
19	3	寶玉 寶釵 知

匯出 選取匯出 就緒 - 總資料數：2666007 上一頁 [1] ~ [200]

7. 同類詞選項

i. 同類詞合併為標籤

同一類的所有詞合併為同一個標籤來統計共現頻率。範例中查詢「主角」，則會將「寶玉」、「鳳姐」這些屬於「主角」類別的詞全部合併計算。

※請注意，若查詢詞不是類別名稱，像是查詢「寶玉」，則會出現無法合併統計的錯誤訊息。若想觀察「寶玉」和「主角」類別中其他詞彙的共現情形，請先至同類詞頁籤中將「寶玉」暫時取消。

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 詞典 停用詞 同類詞 文檔

執行 詞數： 標點 距離限制： 門檻值： 同類詞合併為標籤 ☐ 啟用停用詞 ☐ 忽略

#	Freq.	Co-Words
1	2069	主角 道：
2	1572	」 主角：
3	1436	」 主角 道
4	1200	主角 了，
5	1048	」 主角，
6	688	主角 笑：
7	653	主角 笑 道
8	511	」 主角 笑
9	506	」 主角 了
10	471	主角 見，
11	458	主角 的，
12	432	主角 聽，
13	391	主角 了 聽
14	342	」 主角 聽
15	341	。 主角 了
16	259	主角 來，
17	242	主角 著，
18	242	主角 他，
19	231	主角 立

匯出 選取匯出 就緒 - 總資料數... 上一頁 [1] ~ [2000] 下一頁 設定每頁

同類詞合併為標籤 不啟用同類詞 同類詞合併為標籤 顯示同類詞標籤 只顯示同類詞

啟用	#	Category	Wo
☒	1	主角	[寶玉,
☒	1 (1)	寶玉	
☒	1 (2)	鳳姐	
☒	1 (3)	賈母	
☒	1 (4)	黛玉	
☒	1 (5)	王夫人	
☒	1 (6)	寶釵	
☒	2	配角	[襲人,
☒	2 (1)	襲人	
☒	2 (2)	賈政	

匯入 匯出

ii. 同類詞加上標籤

屬於同類詞的詞彙會掛上所屬的類別名稱，頻率則是各自計算。

第一個範例是以「主角」做為查詢，會列出所有主角類詞彙與所有詞的共現頻率。

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析

執行 詞數： 標點 距離限制： 門檻值： 顯示同類詞標籤 ☐ 啟用停用詞 ☐ 忽略

新增 編輯 刪除

#	Freq.	Co-Words	啟用	#	Category	Wo
1	882	寶玉[主角] 道：	☒	1	主角	[寶玉]
2	681	寶玉[主角] 了：	☒	1 (1)	寶玉	
3	634	寶玉[主角] 了 道	☒	1 (2)	鳳姐	
4	590	寶玉[主角] 了，	☒	1 (3)	賈母	
5	542	寶玉[主角] 了，	☒	1 (4)	黛玉	
6	386	寶玉[主角] 道：	☒	1 (5)	王夫人	
7	317	鳳姐[主角] 道：	☒	1 (6)	寶釵	
8	298	寶玉[主角] 笑：	☒	2	配角	[襲人]
9	289	黛玉[主角] 道：	☒	2 (1)	襲人	
10	282	寶玉[主角] 笑道	☒	2 (2)	賈政	
11	280	寶玉[主角] 了：				
12	254	寶玉[主角] 了 道				
13	244	寶玉[主角] 了 道				
14	239	鳳姐[主角] 了：				
15	238	寶玉[主角] 的，				
16	237	寶玉[主角] 見，				
17	232	黛玉[主角] 了：				
18	228	寶玉[主角] 聽，				

匯出 選取匯出 就緒 - 總資料數... 上一頁 [1] ~ [2000] 下一頁 設定每頁

匯入 匯出

第二個範例是以「寶玉」做為查詢，會列出「寶玉」與所有詞的共現頻率。

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析

執行 詞數： 標點 距離限制： 門檻值： 顯示同類詞標籤 ☐ 啟用停用詞 ☐ 忽略

新增 編輯 刪除

#	Freq.	Co-Words	啟用	#	Category	Wo
43	47	寶玉[主角] 忙，	☒	1	主角	[寶玉]
44	47	寶玉[主角] 忙	☒	1 (1)	寶玉	
45	47	寶玉[主角] 著	☒	1 (2)	鳳姐	
46	45	寶玉[主角] 便	☒	1 (3)	賈母	
47	44	寶玉[主角] 和，	☒	1 (4)	黛玉	
48	44	寶玉[主角] 了	☒	1 (5)	王夫人	
49	43	寶玉[主角] 賈母[主角]，	☒	1 (6)	寶釵	
50	43	寶玉[主角] 又：	☒	2	配角	[襲人]
51	43	寶玉[主角] 襲人[配角]，	☒	2 (1)	襲人	
52	43	寶玉[主角] 要，	☒	2 (2)	賈政	
53	43	寶玉[主角] 上，				
54	42	寶玉[主角] 將，				
55	42	寶玉[主角] 問：				
56	42	寶玉[主角] 寶釵，				
57	41	寶玉[主角] 與，				
58	41	寶玉[主角] 著				
59	40	寶玉[主角] 黛玉[主角]，				
60	40	寶玉[主角] 人，				

匯出 選取匯出 就緒 - 總資料數... 上一頁 [1] ~ [2000] 下一頁 設定每頁

匯入 匯出

iii. 只顯示包含同類詞的組合

查詢結果組合中全部都要出現在同類詞頁籤中，而且要包含查詢詞。

第一個範例是以「主角」做為查詢，查詢結果組合至少有個詞屬「主角」類別。

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析

執行 詞數： 標點 距離限制： 門檻值： 只顯示同類詞 ☐ 啟用停用詞 ☐ 忽略

新增 編輯 刪除

#	Freq.	Co-Words	啟用	#	Category	Wo
1	4	鳳姐[主角] 王夫人[主角]	☒	1	主角	[寶玉,
2	3	寶玉[主角] 王夫人[主角]	☒	1 (1)	寶玉	
3	2	寶玉[主角] 襲人[配角] 賈	☒	1 (2)	鳳姐	
4	1	賈母[主角] 王夫人[主角]	☒	1 (3)	賈母	
5	1	賈母[主角] 王夫人[主角]	☒	1 (4)	黛玉	
6	1	賈母[主角] 王夫人[主角] 黛玉[主...	☒	1 (5)	王夫人	
7	1	黛玉[主角] 王夫人[主角] 襲人[配角]	☒	1 (6)	寶釵	
8	1	鳳姐[主角] 賈母[主角] 黛玉[主角]	☒	2	配角	[襲人,
9	1	寶玉[主角] 賈政[配角] 黛玉[主角]	☒	2 (1)	襲人	
10	1	寶玉[主角] 王夫人[主角] 鳳姐[主角]	☒	2 (2)	賈政	

匯出 選取匯出 就緒 - 總資料數... 上一頁 [1] ~ [2000] 下一頁 設定每頁

匯入 匯出

第二個範例是以“寶玉 寶釵”做為查詢，詞數設為 3，查詢結果組合中「寶玉」、「寶釵」都要出現，第三個詞則要是類別詞。

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析

執行 詞數： 標點 距離限制： 門檻值： 只顯示同類詞 ☐ 啟用停用詞 ☐ 忽略

新增 編輯 刪除

#	Freq.	Co-Words	啟用	#	Category	Wo
1	2	寶玉[主角] 寶釵 黛玉[主角]	☒	1	主角	[寶玉,
2	1	寶玉[主角] 寶釵 賈母[主角]	☒	1 (1)	寶玉	
3	1	寶玉[主角] 寶釵 王夫人[主角]	☒	1 (2)	鳳姐	
4	1	寶玉[主角] 寶釵 鳳姐[主角]	☒	1 (3)	賈母	
			☒	1 (4)	黛玉	
			☒	1 (5)	王夫人	
			☒	1 (6)	寶釵	
			☒	2	配角	[襲人,
			☒	2 (1)	襲人	
			☒	2 (2)	賈政	

匯出 選取匯出 就緒 - 總資料數... 上一頁 [1] ~ [2000] 下一頁 設定每頁

匯入 匯出

十七、 共現分析

功能

1. 輸入搜尋詞並按「執行」(以標點為例)

系統 設定											
標的語料庫		詞頻列表		關鍵詞檢索		搭配詞		關鍵詞分布		共現頻率	
寶玉		標點		執行		不啟用同類詞		☐ 啟用停用詞		☐ 忽略變形	
#	Word.	z_score	t_score	MI	MI^3	chi-sq...	simple-II	dice co...	odds-r...	log-lik...	
1	「	0.963	4.996	4.751	51.521	5.991	0.935	0.001	25.929	2.860	
2	，	1.391	1.412	0.043	0.000	-13429...	-74.116	0.072	26.986	2.046	
3	。	-22.799	-16.253	-0.977	-0.447	124704...	747.925	0.033	21.649	187.798	
4	；	-1.748	-1.055	-1.459	-1.491	-58.876	2.618	0.000	25.910	0.659	
5	：	19.199	29.951	1.283	1.015	304141...	-430.821	0.135	31.776	345.940	
6	？	-13.620	-7.495	-1.723	-2.459	37514...	124.839	0.010	25.367	35.366	
7	！	-15.976	-6.855	-2.441	-6.990	205372...	96.097	0.004	25.527	32.090	
8	—	0.384	0.428	0.308	0.014	1.279	-0.796	0.002	25.928	0.077	
9	《	-1.435	-1.120	-0.715	-0.176	-63.771	4.263	0.002	25.900	0.656	
10	》	-1.435	-1.120	-0.715	-0.176	-63.771	4.263	0.002	25.900	0.656	
11	「	-23.043	-13.438	-1.556	-1.810	-53128...	415.149	0.019	23.983	117.920	
12	」	24.509	42.963	1.620	2.041	973336...	-493.744	0.170	35.074	656.935	
13	『	5.060	2.612	1.444	1.446	2201.0...	20.860	0.005	25.781	7.767	

匯出

2. 可依不同共現指標排序

系統 設定											
標的語料庫		詞頻列表		關鍵詞檢索		搭配詞		關鍵詞分布		共現頻率	
寶玉		標點		執行							
#	Word.	z_sco...	t_score	MI	MI^3	chi-sq...	simple...	dice c...	odds-r...	log-lik...	
1	」	24.509	42.963	1.620	2.041	2149.2...	-493.744	0.170	35.074	656.935	
2	：	19.199	29.951	1.283	1.015	1046.0...	-430.821	0.135	31.776	345.940	
3	道	17.342	31.058	1.681	2.284	1052.3...	-231.234	0.137	29.483	303.783	
4	笑道	11.589	20.945	1.708	2.392	463.306	-100.324	0.089	27.346	130.726	
5	見	11.435	19.090	1.479	1.554	387.866	-124.687	0.091	27.438	117.256	
6	聽說	11.110	20.888	1.822	2.905	458.414	-80.986	0.083	27.175	124.801	
7	聽說	7.030	17.593	2.647	8.906	316.899	-6.105	0.033	26.332	66.422	
8	忙	6.684	10.603	1.331	1.134	116.631	-49.649	0.047	26.445	36.779	
9	甄	5.362	15.581	3.078	14.007	244.731	3.325	0.018	26.147	44.875	
10	因	5.115	7.059	0.929	0.386	51.615	-45.274	0.042	26.310	18.256	
11	”	4.931	9.580	1.916	3.381	92.741	-14.238	0.021	26.151	24.909	

系統 設定											
標的語料庫		詞頻列表		關鍵詞檢索		搭配詞		關鍵詞分布		共現頻率	
寶玉		標點		執行							
#	Word.	z_score	t_score	MI	MI^3	chi-s...	simple...	dice c...	odds-r...	log-lik...	
1	」	24.509	42.963	1.620	2.041	2149.2...	-493.744	0.170	35.074	656.935	
2	道	17.342	31.058	1.681	2.284	1052.3...	-231.234	0.137	29.483	303.783	
3	：	19.199	29.951	1.283	1.015	1046.0...	-430.821	0.135	31.776	345.940	
4	笑道	11.589	20.945	1.708	2.392	463.306	-100.324	0.089	27.346	130.726	
5	聽說	11.110	20.888	1.822	2.905	458.414	-80.986	0.083	27.175	124.801	
6	見	11.435	19.090	1.479	1.554	387.866	-124.687	0.091	27.438	117.256	
7	。	-22.799	-16.253	-0.977	-0.447	378.893	747.925	0.033	21.649	187.798	
8	聽說	7.030	17.593	2.647	8.906	316.899	-6.105	0.033	26.332	66.422	
9	甄	5.362	15.581	3.078	14.007	244.731	3.325	0.018	26.147	44.875	
10	我	-35.424	-14.135	-2.651	-8.950	222.602	403.440	0.008	24.279	144.719	
11	「	-23.043	-13.438	-1.556	-1.810	208.836	415.149	0.019	23.983	117.920	

3. 同類詞及停用詞操作方法同共現頻率

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性 字串頻率 詞語趨勢										
主角 [原搜尋詞：寶玉]		標點	執行		同類詞合併為標籤		☐ 啟用停用詞		☐ 忽略變形	
#	Word.	z_score	t_score	MI	MI^3	chi-squa...	simple-II	dice coe...	odds-rat...	log-likeli...
1	r	0.953	4.373	4.398	40.860	4.532	0.743	0.000	20.079	2.648
2	,	-0.245	-0.244	-0.007	-0.000	4273.978	14.682	0.087	19.950	0.062
3	。	-24.835	-17.915	-0.942	-0.402	-443864...	926.411	0.042	16.800	229.336
4	；	-2.511	-1.340	-1.812	-2.858	-43.975	3.931	0.000	20.065	1.129
5	：	22.519	35.687	1.329	1.127	-628872...	-565.314	0.167	24.963	493.357
6	？	-18.190	-9.106	-1.997	-3.824	93207.584	176.821	0.009	19.608	54.538
7	！	-21.409	-8.128	-2.794	-10.485	1368368...	132.513	0.004	19.751	47.133
8	—	0.764	0.922	0.540	0.075	16.579	-1.795	0.002	20.083	0.342
9	《	-1.317	-1.076	-0.583	-0.095	147.460	4.514	0.003	20.061	0.597
10	》	-1.317	-1.076	-0.583	-0.095	147.460	4.514	0.003	20.061	0.597
11	「	-32.335	-16.801	-1.889	-3.239	1028338...	610.708	0.018	18.408	194.000
12	」	28.066	49.530	1.639	2.115	5533863...	-634.168	0.206	27.432	883.441
13	『	-8.575	-4.600	-1.797	-2.789	52811.959	46.424	0.004	19.951	13.349

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性 字串頻率 詞語趨勢										
寶玉		標點	執行		只顯示同類詞		☐ 啟用停用詞		☐ 忽略變形	
#	Word.	z_score	t_score	MI	MI^3	chi-squa...	simple-II	dice coe...	odds-rat...	log-likeli...
1	黛玉[主角]	1.330	1.476	0.301	0.013	571.515	-9.762	0.020	25.988	0.928

十八、 參照語料庫

功能

1. 使用方法與標的語料庫完全相同。
2. 其中一方重新斷詞時，另一方也會重新斷詞

系統 設定

標的語料庫

詞頻列表

關鍵詞檢索

搭配詞

關鍵詞分布

共現頻率

共現分析

參照語料庫

詞語顯著性

匯入

編輯註釋...

重新斷詞

移除選擇...

#	標題	詞	段落	檔案路徑	
1 081	占旺相四美釣...	5271	12	E:\Workspace\紅樓夢 全集\081	占旺相四美釣游魚 奉...
2 082	老學究講義警...	6541	12	E:\Workspace\紅樓夢 全集\082	老學究講義警頑心 病...
3 083	省宮闈賈元妃...	6325	16	E:\Workspace\紅樓夢 全集\083	省宮闈賈元妃染恙 鬧...
4 084	試文字寶玉始...	5968	16	E:\Workspace\紅樓夢 全集\084	試文字寶玉始提親 探...
5 085	賈存周報升郎...	6309	21	E:\Workspace\紅樓夢 全集\085	賈存周報升郎中任 薛...
6 086	受私賄老官翻...	5299	14	E:\Workspace\紅樓夢 全集\086	受私賄老官翻案牘 寄...
7 087	感深秋撫琴悲...	5268	24	E:\Workspace\紅樓夢 全集\087	感深秋撫琴悲往事 坐...
8 088	博庭歡寶玉贊...	5231	11	E:\Workspace\紅樓夢 全集\088	博庭歡寶玉贊孤兒 正...
9 089	人亡物在公子...	4758	15	E:\Workspace\紅樓夢 全集\089	人亡物在公子填詞 蛇...

就緒

移除所有文件

十九、 詞語顯著性

功能

1. 當匯入參照語料庫之後，詞語顯著性資料會立刻計算出來

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性

keyness: log-likelih... 正負向皆顯示 ☐ 啟用停用詞 ☐ 啟用同類詞

#	詞	詞頻(標的)	% (標的)	詞頻(參照)	% (參照)	keyness	p-Value
1	"	0	0	326	0.151	328.010453943...	0.000000000000
2	"	0	0	316	0.146	317.948783577...	0.000000000000
3	笑道	2109	0.447	191	0.088	310.899078262...	0.000000000000
4	忙	1197	0.254	121	0.056	162.466285321...	0.000000000000
5	政	289	0.061	485	0.224	138.389623013...	0.000000000000
6	麼	66	0.014	242	0.112	126.096039361...	0.000000000000
7	鵬	132	0.028	305	0.141	117.563432568...	0.000000000000
8	老爺	205	0.043	371	0.172	114.680435365...	0.000000000000
9	賈	1932	0.409	1522	0.704	105.477604169...	0.000000000000
10	紫	197	0.042	347	0.161	104.302631437...	0.000000000000
11	叫	1041	0.221	956	0.442	101.914248800...	0.000000000000

2. keyness 可選 log-likelihood 或 Chi-square

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性

keyness: log-likelihood 正負向皆顯示 ☐ 啟用停用詞 ☐ 啟用同類詞

#	詞	詞頻(標的)	% (標的)	詞頻(參照)	% (參照)	keyness	p-Value
1	Chi-square	0	0	326	0.151	328.010453943...	0.000000000000
2		0	0	316	0.146	317.948783577...	0.000000000000
3	笑道	2109	0.447	191	0.088	310.899078262...	0.000000000000
4	忙	1197	0.254	121	0.056	162.466285321...	0.000000000000
5	政	289	0.061	485	0.224	138.389623013...	0.000000000000
6	麼	66	0.014	242	0.112	126.096039361...	0.000000000000
7	鵬	132	0.028	305	0.141	117.563432568...	0.000000000000
8	老爺	205	0.043	371	0.172	114.680435365...	0.000000000000
9	賈	1932	0.409	1522	0.704	105.477604169...	0.000000000000
10	紫	197	0.042	347	0.161	104.302631437...	0.000000000000
11	叫	1041	0.221	956	0.442	101.914248800...	0.000000000000

3. 正向顯著會以黑字呈現，負向會以紅字呈現。可選是否只顯示正向或負向顯著詞

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性

keyness: log-likelih... 只顯示正向 ☐ 啟用停用詞 ☐ 啟用同類詞

#	詞	詞頻(標的)	% (標的)	詞頻(參照)	% (參照)	keyness	p-Value
1	笑道	2109	0.447	191	0.088	310.899078262...	0.000000000000
2	忙	1197	0.254	121	0.056	162.466285321...	0.000000000000
3	之	1493	0.316	277	0.128	100.477479628...	0.000000000000
4	吃	1036	0.219	160	0.074	91.3214972094...	0.000000000000
5	皆	365	0.077	12	0.006	85.3657635568...	0.000000000000
6	方	592	0.125	63	0.029	77.0433419775...	0.000000000000
7	林黛玉	267	0.057	8	0.004	64.0225121268...	0.000000000000
8	尤	458	0.097	49	0.023	59.3376974953...	0.000000000000
9	丫	228	0.048	6	0.003	56.4335413579...	0.000000000000
10	襲	231	0.049	7	0.003	55.2281192841...	0.000000000000
11	亦	351	0.074	31	0.014	52.6610298427...	0.000000000000

系統 設定							
標的語料庫		詞頻列表	關鍵詞檢索	搭配詞	關鍵詞分布	共現頻率	共現分析
keyness: log-likelih...		只顯示負向	☐ 啟用停用詞		☐ 啟用同類詞		
#	詞	詞頻(標的)	% (標的)	詞頻(參照)	% (參照)	keyness	p-Value
1	"	0	0	326	0.151	328.010453943...	0.000000000000
2	"	0	0	316	0.146	317.948783577...	0.000000000000
3	政	289	0.061	485	0.224	138.389623013...	0.000000000000
4	麼	66	0.014	242	0.112	126.096039361...	0.000000000000
5	鵬	132	0.028	305	0.141	117.563432568...	0.000000000000
6	老爺	205	0.043	371	0.172	114.680435365...	0.000000000000
7	賈	1932	0.409	1522	0.704	105.477604169...	0.000000000000
8	紫	197	0.042	347	0.161	104.302631437...	0.000000000000
9	叫	1041	0.221	956	0.442	101.914248800...	0.000000000000
10	沒有	317	0.067	437	0.202	97.8380434593...	0.000000000000
11	里	543	0.115	590	0.273	90.0970301369...	0.000000000000

4. 啟用停用詞時，停用詞不會出現在列表中

系統 設定							
標的語料庫		詞頻列表	關鍵詞檢索	搭配詞	關鍵詞分布	共現頻率	共現分析
keyness: log-likelih...		正負向皆顯示	☒ 啟用停用詞		☐ 啟用同類詞		
#	詞	詞頻(標的)	% (標的)	詞頻(參照)	% (參照)	keyness	p-Value
1	笑道	2109	0.447	191	0.088	310.899078262...	0.000000000000
2	忙	1197	0.254	121	0.056	162.466285321...	0.000000000000
3	政	289	0.061	485	0.224	138.389623013...	0.000000000000
4	麼	66	0.014	242	0.112	126.096039361...	0.000000000000
5	鵬	132	0.028	305	0.141	117.563432568...	0.000000000000
6	老爺	205	0.043	371	0.172	114.680435365...	0.000000000000
7	賈	1932	0.409	1522	0.704	105.477604169...	0.000000000000
8	紫	197	0.042	347	0.161	104.302631437...	0.000000000000
9	叫	1041	0.221	956	0.442	101.914248800...	0.000000000000
10	之	1493	0.316	277	0.128	100.477479628...	0.000000000000
11	沒有	317	0.067	437	0.202	97.8380434593...	0.000000000000

5. 啟用同類詞時，所有同類詞合併為一個詞(以其類別標籤做代表)，去計算顯著性

系統 設定							
關鍵詞檢索		搭配詞	關鍵詞分布	共現頻率	共現分析	參照語料庫	詞語顯著性
keyness: log-l...		正負向皆顯示	☒ 啟用停用詞		☒ 啟用同類詞		
#	詞	詞頻(標的)	% (標的)	詞頻(參照)	% (參照)	keyne...	p-Val...
1	主角	4400	0.932	2319	1.073	12.9247...	0.0003...
2	又	3771	0.799	1438	0.666	15.4690...	0.0000...
3	這	3488	0.739	1259	0.583	23.4793...	0.0000...
4	來	3395	0.719	1375	0.636	6.46719...	0.0109...
5	道	3209	0.68	2133	0.987	75.1549...	0.0000...
6	著	3175	0.673	1877	0.869	32.7430...	0.0000...
7	人	2899	0.614	1315	0.609	0.03186...	0.8583...
8	兒	2896	0.614	1388	0.642	0.85987...	0.3537...
9	個	2731	0.579	924	0.428	28.7072...	0.0000...
10	去	2660	0.564	1137	0.526	1.63327...	0.2012...
11	有	2388	0.506	909	0.421	9.97780...	0.0015...

詞典 停用詞 同類詞 文檔檢視 詮釋資料欄位				
新增		編輯		刪除
#	啟用	Category	Words	Source
1	☒	▼ 主角	[寶玉, 賈...	本次新增
2	☒	寶玉		本次新增
3	☒	賈母		本次新增
4	☒	黛玉		本次新增

二十、 字串頻率

功能

- 輸入字數 N 並按「搜尋」。(以 5+文章為例)

系統 設定

關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性 字串頻率 詞語趨勢

搜尋： 搜尋 N： 邊界： 門檻值：

#	字串	頻率
1	笑道：「你	315
2	玉笑道：「	311
3	寶玉道：「	287
4	。」說著，	282
5	笑道：「我	277
6	笑道：「這	250
7	」寶玉道：	216
8	寶玉笑道：	214
9	，笑道：「	188

匯出

2. 可設門檻值

系統 設定

關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性 字串頻率 詞語趨勢

搜尋： 搜尋 N： 邊界： 門檻值：

#	字串	頻率
1	笑道：「你	315
2	玉笑道：「	311
3	寶玉道：「	287
4	。」說著，	282
5	笑道：「我	277
6	笑道：「這	250
7	」寶玉道：	216
8	寶玉笑道：	214

匯出

3. 輸入搜尋字串，可查找包含有查詢的字串

系統 設定

關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性 字串頻率 詞語趨勢

搜尋： 搜尋 N： 邊界： 門檻值：

#	字串	頻率
1	黛玉道：「	103
2	黛玉笑道：	86
3	」黛玉道：	60
4	。」林黛玉	56
5	」黛玉笑道	54
6	。」黛玉道	47
7	。」黛玉笑	43
8	了。」黛玉	34
9	林黛玉道：	27

匯出

4. 查詢兩個以上的詞、用空白隔開表示出現任何查詢的字串

系統 設定

關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性 字串頻率 詞語趨勢

搜尋： "黛玉 寶玉" 搜尋 N： 5 邊界： 文章 門檻值：

#	字串	頻率
1	黛玉見寶玉	5
2	黛玉，寶玉	4
3	寶玉林黛玉	3
4	黛玉，寶玉	2
5	寶玉和黛玉	2
6	寶玉，黛玉	2
7	黛玉和寶玉	2
8	黛玉被寶玉	1
9	寶玉見黛玉	1

匯出

系統 設定

關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性 字串頻率 詞語趨勢

搜尋： 黛玉 寶玉 搜尋 N： 5 邊界： 文章 門檻值：

#	字串	頻率
1	寶玉道：「	287
2	」寶玉道：	216
3	寶玉笑道：	214
4	。」「寶玉道	170
5	」寶玉笑道	166
6	了。」「寶玉	165
7	。」「寶玉聽	140
8	。」「寶玉笑	129
9	，寶玉聽了	109

匯出

- 二十一、 更正斷詞列表
功能
1. 新增更正斷詞詞組

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析

搜尋： 執行 2 文章 ☐ 啟用停用詞

#	Fr...	N-gram
13	48	， 是
14	44	這 個
15	43	我 們 要
16	42	， 而
17	41	」，
18	38	新 政 府
19	37	， 這
20	37	。 這
21	36	人 民 的
22	36	國 家 的
23	36	， 來
24	35	台 灣 的
25	24	。 為

文檔檢視 詮釋資料欄位 更正斷詞列表

原始： 新政府 更正： 新政府 新增 刪除

原始組合	更正組合	來源
		新增

匯入 匯出新... 匯出完...

系統 設定

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析

搜尋： 執行 2 文章 ☐ 啟用停用詞

#	Fr...	N-gram
13	48	， 是
14	44	這 個
15	43	我 們 要
16	42	， 而
17	41	」，
18	38	新 政 府
19	37	， 這
20	37	。 這
21	36	人 民 的
22	36	國 家 的
23	36	， 來
24	35	台 灣 的
25	24	。 為

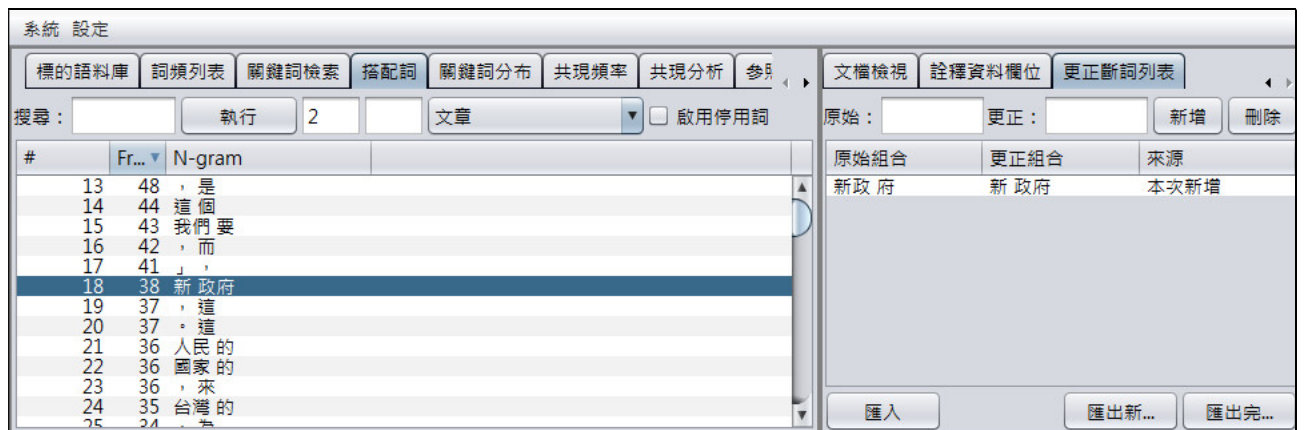
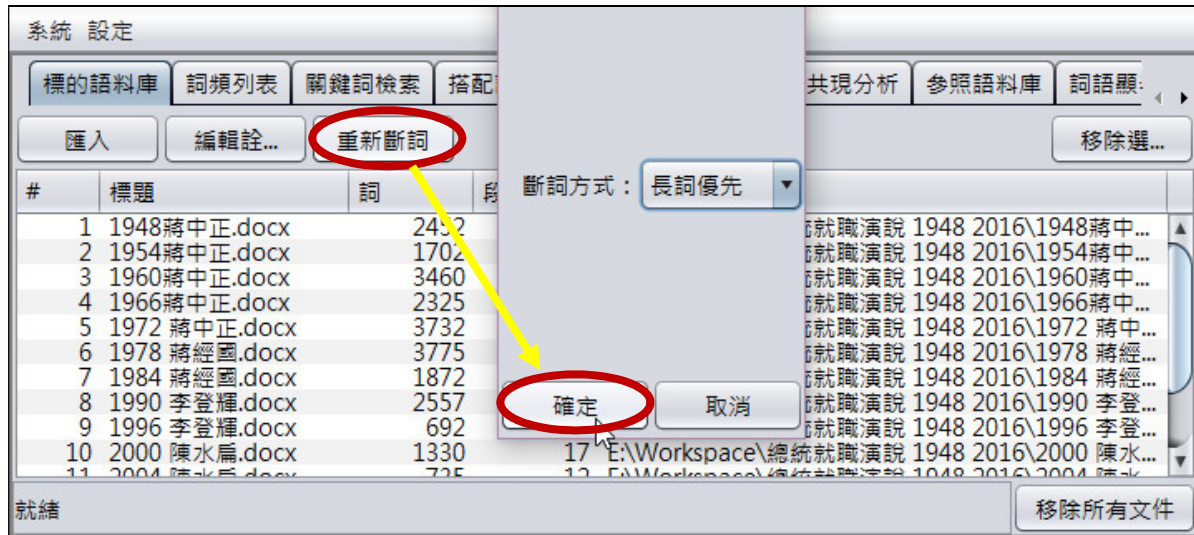
文檔檢視 詮釋資料欄位 更正斷詞列表

原始： 更正： 新增 刪除

原始組合	更正組合	來源
新政府	新政府	本次新增

匯入 匯出新... 匯出完...

2. 選擇重新斷詞以更正詞組



3. 可以匯入匯出列表，方法同辭典、停用詞、同類詞

1. 尋找變形組合，找出建議的變形組合



2. 新增編輯及刪除變形等操作方法同同類詞

3. 詞頻列表與忽略變形



4. 關鍵詞檢索與忽略變形

標的語料庫		詞頻列表		關鍵詞檢索		搭配詞		關鍵詞分布		共現頻率		共現分析		參照語料庫		詞語顯著性		字串頻率		詞語趨勢				
搜尋詞：abandon				執行		左數：2		右數：2		文章		☐ 同類詞 ☐ 停用詞 ☒ 忽略變形												
#	左方出現詞			關鍵詞		右方出現詞			來源文章															
13	and must			abandon		the idea			UK_LLC_2010_v25n2_7.txt															
14	researchers who			abandon		universal ,			UK_LLC_2010_v25n3_1.txt															
15	us to			abandon		the popular			UK_LLC_2010_v25n3_2.txt															
16	be to			abandon		a strong			US_DHQ_2010_4.1_1.txt															
17	development .			Abandoning		the restrictive			CA_DS_2009_v1n2_4.txt															
18	but not			abandoning		) the			UK_LLC_2009_v24n1_8.txt															
19	we are			abandoning		the goal			US_DHQ_2009_3.2_1.txt															
20	passable without			abandoning		specific bad			US_DHQ_2009_3.3_4.txt															
21	; by			abandoning		your virtue			UK_LLC_2010_v25n2_8.txt															
22	ena without			abandoning		the rich			UK_LLC_2010_v25n4_4.txt															
23	so we			abandoned		Terence .			CA_DS_2009_v1n3_4.txt															
24	they have			abandoned		for ages			TW_DADH_2009_2.txt															
匯出		搜尋到40筆結果																			上一頁		下一頁	

5. 搭配詞與忽略變形

標的語料庫		詞頻列表		關鍵詞檢索		搭配詞		關鍵詞分布		共現頻率		共現分析		參照語料庫		詞語顯著性		字串頻率		詞語趨勢	
搜尋：		執行		2				文章				☐ 啟用停用詞				☐ 忽略變形					
#		Freq.		▼		N-gram															
1		29566				..															
2		11488				of the															
3		7087				, and															
4		6308				in the															
5		5567				. The															
6		5523				).															
7		4408				, the															
8		3539				to the															
9		2641				),															
10		2624				and the															
11		2472				. In															
12		2444				(															
匯出																					

標的語料庫		詞頻列表		關鍵詞檢索		搭配詞		關鍵詞分布		共現頻率		共現分析		參照語料庫		詞語顯著性		字串頻率		詞語趨勢	
搜尋：				執行		2				文章				☐ 啟用停用詞		☒ 忽略變形					
#		Freq.		▼		N-gram															
1		29566				..															
2		11623				of the															
3		7110				, and															
4		6879				in the															
5		5694				. the															
6		5523				).															
7		4566				, the															
8		3569				to the															
9		2688				and the															
10		2649				, a															
11		2641				),															
12		2528				in															
匯出																					

6. 關鍵詞分布與忽略變形

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性 字串頻率									
搜尋詞：abandon		搜尋	☐ 啟用同... ☐ 啟用停... ☐ 忽略變形						
分布圖	搜尋詞	出現次數	文章總詞數	檔案來源					
	abandon	2	7346	CA_DS_...					
	abandon	1	9336	CA_DS_...					
	abandon	1	12351	UK_LLC_...					
	abandon	1	15325	UK_LLC_...					
	abandon	1	22624	US_DH_...					
	abandon	1	8394	US_DH_...					
	abandon	1	10927	US_DH_...					
	abandon	1	10635	US_DH_...					
	abandon	2	13965	US_DH_...					
	abandon	1	9752	UK_LLC_...					
	abandon	1	11614	UK_LLC_...					
	abandon	1	12109	UK_LLC_...					
	abandon	1	11632	UK_LLC_...					
	abandon	1	25498	US_DH_...					

標的語料庫	詞頻列表	關鍵詞檢索	搭配詞	關鍵詞分布	共現頻率	共現分析	參照語料庫	詞語顯著性	字串
搜尋詞：abandon		搜尋		☐ 啟用同... ☐ 啟用停... ☒ 忽略變形					
分布圖					搜尋詞	出現次數	文章總詞數	檔案來源	
					abandon	2	7346	CA_DS_...	
					abandon	1	10343	CA_DS_...	
					abandon	1	9336	CA_DS_...	
					abandon	1	5788	CA_DS_...	
					abandon	1	5495	TW_DA...	
					abandon	1	7760	TW_DA...	
					abandon	1	9063	UK_IJH...	
					abandon	1	7384	UK_LLC...	
					abandon	1	12351	UK_LLC...	
					abandon	1	11552	UK_LLC...	
					abandon	1	12255	UK_LLC...	
					abandon	1	15325	UK_LLC...	
					abandon	2	11016	UK_LLC...	
					abandon	1	22624	US_DH...	

7. 共現頻率與忽略變形

標的語料庫

詞頻列表

關鍵詞檢索

搭配詞

關鍵詞分布

共現頻率

共現分析

參照語料庫

詞語顯著性

字串

搜尋詞：

lon

執行

詞數：

2

標點

距離限制：

門檻值：

不啟用同類詞

☐ 啟用停用詞

☐ 忽略變形

#	Freq.	Co-Words
1	8	abandon the
2	7	, abandon
3	7	abandon to
4	5	. abandon
5	4	abandon of
6	4	abandon that
7	4	abandon we
8	3	a abandon
9	3	abandon and
10	3	abandon in
11	2	This abandon
12	2	abandon as

匯出

就緒

匯出 就緒

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性 字串

搜尋詞: lon 執行 詞數: 2 標點 距離限制: 門檻值: 不啟用同類詞 啟用停用詞 忽略變形

#	Freq.	Co-Words
1	19	abandon the
2	15	, abandon
3	15	. abandon
4	11	abandon be
5	11	abandon that
6	11	abandon to
7	10	abandon of
8	9	a abandon
9	9	abandon in
10	8	abandon we
11	6	abandon and
12	5	abandon not

匯出 就緒

匯出 就緒

8. 共現分析與忽略變形

標的語料庫		詞頻列表		關鍵詞檢索		搭配詞		關鍵詞分布		共現頻率		共現分析		參照語料庫		詞語顯著性		字串	
bandon		標點		執行		不啟用同類詞		☐ 啟用停用詞		☐ 忽略變形									
#	Word.	z_score	t_score	MI	MI^3	chi-squ...	simple-II	dice co...	odds-r...	log-lik...									
1	,	0.865	1.055	0.571	0.090	-13839...	-2.171	0.000	19550.8...	0.637									
2	.	0.221	0.233	0.150	0.002	-0.002	-0.537	0.000	16443.9...	0.032									
3	a	0.995	1.524	1.232	0.898	-66.555	-1.220	0.000	17941.8...	0.820									
4	affiliati...	0.999	32.036	10.006	481.322	256.103	4.026	0.063	16981.3...	5.212									
5	also	0.872	2.432	2.961	12.472	31.195	0.040	0.001	16846.0...	1.047									
6	an	0.705	1.297	1.760	2.620	-2.905	-0.350	0.000	16669.0...	0.462									
7	and	0.604	0.749	0.619	0.114	-11.385	-0.975	0.000	17117.5...	0.244									
8	appear	0.985	7.940	6.023	104.978	15.288	1.657	0.007	16966.0...	2.800									
9	as	1.048	2.058	1.948	3.550	44.081	-0.618	0.000	17565.3...	1.122									
10	attitudes	0.999	26.708	9.482	409.651	177.860	3.712	0.051	16980.8...	4.888									
11	avatar	1.410	27.119	8.530	298.224	413.106	6.282	0.038	18147.4...	8.671									
12	be	0.567	0.861	1.206	0.843	-8.769	-0.407	0.000	16522.3	0.243									
匯出																			

匯出

標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性 字串頻率										
bandon		標點	執行		不啟用同類詞		☐ 啟用停用詞		☒ 忽略變形	
#	Word.	z_score	t_score	MI	MI^3	chi-squ...	simple-II	dice co...	odds-r...	log-lik...
1	,	0.832	0.939	0.349	0.020	-5553.1...	-3.295	0.000	7283.039	0.519
2	.	0.965	1.113	0.413	0.034	-781.398	-3.741	0.000	7415.353	0.708
3	's	0.330	0.403	0.578	0.093	-2.428	-0.312	0.000	6551.759	0.063
4	(	0.513	0.611	0.506	0.062	38.036	-0.862	0.000	6648.173	0.152
5	)	-0.110	-0.106	-0.108	-0.001	-8.588	0.180	0.000	6468.091	0.005
6	1767	0.999	30.629	9.877	462.905	234.079	3.949	0.043	6663.177	5.155
7	a	1.360	1.839	0.871	0.318	643.226	-3.439	0.001	7328.016	1.392
8	accessi...	0.976	6.363	5.408	75.990	9.645	1.303	0.010	6659.430	2.421
9	affiliati...	0.996	16.163	8.040	249.723	64.830	2.848	0.031	6662.722	4.004
10	age	0.971	5.668	5.092	63.420	7.561	1.124	0.009	6658.468	2.235
11	almost	0.961	4.894	4.695	49.736	-9.111	0.904	0.007	6656.924	2.004
12	also	0.671	1.170	1.605	1.985	1.565	-0.376	0.001	6608.570	0.388

匯出

9. 詞語顯著性與忽略變形

關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性 字串頻率 詞語趨勢										
keyness: log-lik...		正負向皆顯示		☐ 啟用停用詞		☐ 啟用同類詞		☐ 忽略變形		
#	詞	詞頻(標的)	%(標的)	詞頻(參照)	%(參照)	keyness	p-Value			
1	2011	12	0.001	1890	0.077	704.387084982162	0.000000000000			
2	2012	3	0	1216	0.049	471.510988071626	0.000000000000			
3	Shakespeare	889	0.061	141	0.006	468.045185738935	0.000000000000			
4	2013	0	0	1031	0.042	413.055020280042	0.000000000000			
5	TEI	1069	0.074	395	0.016	341.242996276211	0.000000000000			
6	=	514	0.036	76	0.003	278.084186211311	0.000000000000			
7	h	884	0.061	416	0.017	223.274193325883	0.000000000000			
8	q	367	0.025	49	0.002	206.017137240386	0.000000000000			
9	l	1781	0.123	5148	0.208	172.493905096152	0.000000000000			
10	24	594	0.041	258	0.01	163.217369626885	0.000000000000			
11	markup	500	0.035	181	0.007	162.399134707231	0.000000000000			
12	Bentham	0	0	400	0.016	160.254130079551	0.000000000000			

匯出

關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 參照語料庫 詞語顯著性 字串頻率 詞語趨勢										
keyness: log-lik...		正負向皆顯示		☐ 啟用停用詞		☐ 啟用同類詞		☒ 忽略變形		
#	詞	詞頻(標的)	%(標的)	詞頻(參照)	%(參照)	keyness	p-Value			
1	2011	12	0.001	1890	0.077	704.387084982162	0.000000000000			
2	2012	3	0	1216	0.049	471.510988071626	0.000000000000			
3	shakespeare	892	0.062	142	0.006	468.929984196333	0.000000000000			
4	2013	0	0	1031	0.042	413.055020280042	0.000000000000			
5	tei	1073	0.074	396	0.016	342.872466394073	0.000000000000			
6	=	514	0.036	76	0.003	278.084186211311	0.000000000000			
7	edition	1667	0.115	1100	0.045	267.175151555903	0.000000000000			
8	le	734	0.051	305	0.012	210.666485086701	0.000000000000			
9	markup	613	0.042	218	0.009	202.054105264341	0.000000000000			
10	dialect	25	0.002	655	0.027	190.996163089260	0.000000000000			
11	encode	953	0.066	603	0.024	163.380912749660	0.000000000000			
12	24	595	0.041	259	0.01	163.132195111765	0.000000000000			

匯出

10. 詞語趨勢與忽略變形

