

CLiC 1.6 User Guide

1. Introduction	1
2. The CLiC corpora	1
3. The CLiC analysis tabs	3
3.1. <i>Functions common to all tabs</i>	4
3.2. <i>Concordance</i>	4
• Concordance – ‘Search the corpora’	4
• Concordance – ‘Only in subsets’	4
• Concordance – ‘Search for terms’	4
• Concordance – ‘Whole phrase’ or ‘Any word’	5
• Concordance – Co-text	5
• Concordance – ‘Results’	5
• Concordance – Basic sorting	6
• Concordance – ‘KWICGrouper’	7
• Concordance – ‘Manage tag columns’	11
3.3. <i>‘Subsets’</i>	13
• Subsets – Results	13
3.4. <i>Clusters</i>	14
3.5. <i>Keywords</i>	14
4. Appendix: List of texts available in CLiC	17

1. Introduction

With the 1.6 release of CLiC in the autumn of 2017 the CLiC interface has changed completely.¹ The tabs for the analysis functions ('Concordance', 'Subsets', 'Keywords', 'Clusters') are now located on the side panel on the right (see Figure 1). The sidebar menu can be toggled by clicking on the menu icon (☰) in the top right corner in order to make more space for the analysis screen.

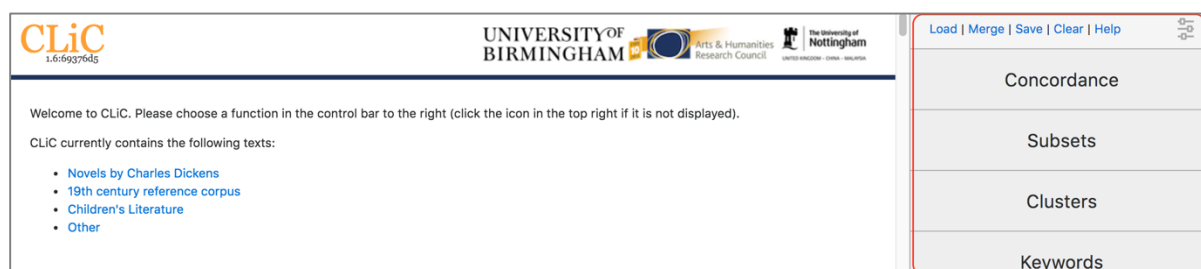


Figure 1: The CLiC home screen with the menu in the side panel on the right.

2. The CLiC corpora

The CLiC homepage, part of which is shown in Figure 1, acts as a table of contents for all the books and their chapters; when clicking onto any of the chapter links you are taken to the full text of that section. As of version 1.6 [released in November 2017], CLiC contains over 130 texts, mainly from the 19th century; for a full list of titles please refer to the appendix. The procedure followed for retrieving, cleaning and importing the most recent texts is described in detail on our Github page.² The texts can be selected individually and combined freely for analysis in any of the CLiC tools.³ You can also choose from one of our three pre-selected corpora: *Novels by Charles Dickens*, *19th Century Reference Corpus* and the *Children's Literature Corpus*. In addition, there is a loosely defined collection of 'Other' texts which include additional GCSE and A-Level titles.

In any of the tools on the sidebar menu you can select any or all of the texts by picking the corpora from a drop-down list or typing their names into a textbox. For example, in the Concordance tool, once you have clicked on the Concordance tab, a textbox labeled 'Search the corpora' will appear (for more information on the Concordance tab, see Section 2.2), as illustrated in Figure 2.

¹ We are grateful to colleagues who have commented on previous drafts of this guide: Jamie Lentin, Anna Cermakova and Lorenzo Mastropiero. Yet, given the new interface, this user guide is still a work in progress. Please get in touch via clic@contacts.birmingham.ac.uk if you have further questions or suggestions for improvement.

² <https://github.com/birmingham-ccr/corpora/blob/master/README.pdf>

³ Please note that we do not recommend selecting all CLiC texts at once, because the tool will not be able to handle this quantity of data and therefore will not show any results. Moreover, any aspect of principled design would be lost when all texts were to be combined; it is better to choose one of our pre-selected corpora or create your own selection.

Figure 2: Selecting corpora in the Concordance tab (same procedure in Subsets and Clusters; for Keywords see Section 2.5).

You can select the pre-selected corpora in their entirety or you can pick individual books from them (again from the drop-down or by typing the title), effectively creating your own subcorpus/subcorpora. For example, you could select several books from Dickens, several books from the *19th Century Novels Reference Corpus* and several books from the *Children's Literature Corpus*. You could also create an author-based corpus, for example by selecting all of Jane Austen's novels by typing "Austen" into the textbox; this will pull out the novels from any of the corpora (see Figure 3).

Figure 3: Example of creating a (temporary) author-based corpus: selecting all of Jane Austen's novels.

The CLiC corpora have been marked up to distinguish between several textual subsets of novels (see the example 1 from *Great Expectations* below).

- **quotes:** any text listed in quotes, i.e. mostly character speech but also thoughts or songs that might appear in quotes
- non-quotes: narration
 - and a special case of non-quotes, **suspensions**, which represent narratorial interruptions of character speech that do not end with sentence-final punctuation. Suspensions are further divided by length:
- - short suspensions have a length up to four words
 - long suspensions have a length of five or more words

- (1) "And on what evidence, Pip," asked Mr. Jaggers, very coolly, as he paused with his handkerchief half way to his nose, "does Provis make this claim?"
"He does not make it," said I, "and has never made it, and has no knowledge or belief that his daughter is in existence."
For once, the powerful pocket-handkerchief failed. My reply was so unexpected that Mr. Jaggers put the handkerchief back into his pocket without completing the usual performance, folded his arms, and looked with stern attention at me, though with an immovable face.

A more detailed discussion of the subsets in CLiC – and the rationale behind this division – can be found in this open access article by Mahlberg et al. (2016).⁴

3. The CLiC analysis tabs

The homepage shows the table of contents of the books in CLiC. Click onto one of the tabs in the side bar to start your analysis. The CLiC logo will take you back to the homepage if you want to see the content page again at a later point.

The CLiC functions can be divided into two groups:

- The ‘**concordance**’ and ‘**subsets**’ tabs both display text (patterns) from the selected books in context. This is where you can analyse the use of particular words and phrases.
- The ‘**cluster**’ and ‘**keywords**’ tabs both show lists of frequent patterns (without context), but they have different applications. In the keywords tab, you can compare the frequency of words and phrases in one corpus with another; the tool will provide a list of those items that are significantly ‘overused’ in the first corpus (for more information, see the ‘Keywords’ section).
-

The CLiC analysis is based on ‘white-space tokenisation’. This means that any sequence of letters that is not interrupted by a white space is considered a word. Some special cases arise, however:

- As mentioned in 2.1 below, the ‘filter’ function in the concordance and subset tabs works differently from the other search functions. The filter does not follow the tokenisation procedure but simply filters for character sequences, i.e. also punctuation. This means that you can filter for round brackets, colons etc. if this is useful for your research.
- Apostrophes: *Oliver* and *Oliver’s* count as the same type in CLiC. Therefore, when you search for *Oliver* in *Oliver Twist* you will retrieve all instances of *Oliver* and *Oliver’s* (and vice versa; 826 results). Note, however, that the filter searches by word form: So if you want to find only *Oliver’s*, for example, you can filter the 826 rows and retrieve 100 entries of *Oliver’s*.
- Two hyphens separate words: for example, *Char--lotte* in *Oliver Twist* (OT.c6.p20) “Oliver’s gone mad! Char--lotte!” counts as two tokens.

⁴ Mahlberg, M., Stockwell, P., de Joode, J., Smith, C., & O’Donnell, M. B. (2016). CLiC Dickens: novel uses of concordances for the integration of corpus stylistics and cognitive poetics. *Corpora*, 11(3), 433–463 [Open access, available from <https://doi.org/10.3366/cor.2016.0102>]

3.1. Functions common to all tabs

- At any point, you can close the menu on the right by clicking on the settings icon in the top right corner (see Figure 4).

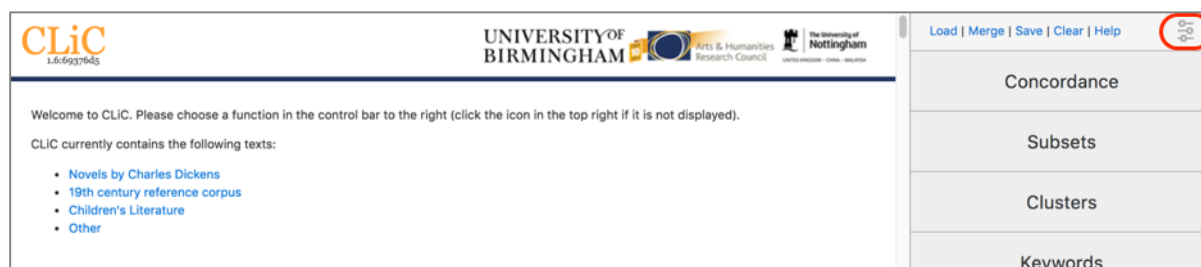


Figure 4: Close the sidebar menu by clicking on the menu icon in the top right corner.

- The buttons in the top row apply to all analysis tabs:
 - ‘Load’**: you can upload a previously exported CLiC csv file to restore your settings and your tag annotation (see the ‘tagging’ section below). Note that the ‘Load’ function will replace any tags on the website with those from the file.
 - ‘Merge’**: The ‘Merge’ function will add the tags from the CSV file to any pre-existing tags. You can also use this function when you have more than one CSV (for example with annotations from several researchers) so you can merge these in order to check to what extent the tag sets overlap or differ.
 - ‘Save’**: save your results, settings and annotated tags in a CSV file.
 - ‘Clear’** will reset the search settings and any tag columns.

For both ‘Load’ and ‘Save’ the results/queries have to be compatible, i.e. they have to be based on the same node word.

3.2. Concordance

Clicking onto the ‘Concordance’ tab will take you to the concordance view. In order to create a concordance, you will need to select a corpus to search in (see the section on ‘The Corpora’ above).

- Concordance – ‘Search the corpora’**

This is where you select a corpus to search in (cf. Section 2). The selection is very flexible and lets you pick a pre-defined corpus (*19th Century Novels Reference Corpus*, *Dickens’s Novels* or *Children’s Literature*) or choose your own subcorpus – either from books from only one of these corpora or combining books across the pre-defined corpora.

- Concordance – ‘Only in subsets’**

Here you can decide whether you want to search through ‘all text’ – the whole book(s) – or just one of the subsets: ‘short suspensions’, ‘long suspensions’, ‘quotes’ and ‘non-quotes’ (cf. Section 2).

- Concordance – ‘Search for terms’**

This is the fundamental parameter of the concordance search – it lets you determine the node word or phrase that forms the basis of the concordance. When you type your search word(s), keep in mind the notes from the tokenisation section above. The node has to be a

valid token according to the white-space tokenisation: for example, a search for a punctuation sign will not retrieve any results. If your research focuses on punctuation markers you can evade this issue by using the filter function in the subset tab: Go to the subset tab, selecting the relevant subset, for example non-quotes, and filtering the rows to the punctuation marker of interest.

- **Concordance – ‘Whole phrase’ or ‘Any word’**

When you have entered several terms, you need to specify whether it is to be searched as one phrase (equivalent to using double quotes in a search engine, e.g. *dense fog*) or any of the words individually (*dense* and *fog*).

- **Concordance – Co-text**

The maximum number of words in the co-text is set at 10 on either side in a concordance (depending on the length of the words and the size of the screen you might see fewer). You can see the full chapter view by clicking on “in bk.” (in book) button at the end of any row.

The screenshot shows the CLiC interface. At the top, there are logos for CLiC 1.6, the University of Birmingham, and the University of Nottingham. Below the logos, it says "Showing 1 to 50 of 92 entries from 1 book". The main table has four columns: "Left", "Node", "Right", and "In bk.". The "In bk." column contains a button labeled "in bk.". To the right of the table is a sidebar with the title "Concordance". It contains a search box labeled "Search the corpora:" with the text "A Christmas Carol: A Ghost Story of Christmas (Charles Dickens)". Below this is a section "Only in subsets:" with a dropdown menu set to "All text". At the bottom of the sidebar is a search box labeled "Search for terms:" with the text "ghost".

Figure 5: The “in bk.” (in book) button leads to the chapter view of the occurrence.

The screenshot shows the CLiC interface displaying the full chapter view of the preface of "A Christmas Carol: A Ghost Story of Christmas, by Charles Dickens". The title is centered at the top. Below the title is the word "PREFACE." followed by a paragraph of text. The text is highlighted in yellow. The paragraph is: "I HAVE endeavoured in this Ghostly little book, to raise the Ghost of an Idea, which shall not put my readers out of humour with themselves, with each other, with the season, or with me. May it haunt their houses pleasantly, and no one wish to lay it." Below this paragraph is a line of text: "Their faithful Friend and Servant, C. D. December, 1843." The text is highlighted in yellow. The interface also shows the CLiC logo and the University of Birmingham and Nottingham logos at the top.

Figure 6: The “in bk.” view shows the whole chapter – in the case of this preface it is a very short chapter. (Note that all authorial text occurring before the official first chapter, is counted as ‘chapter 0’ in CLiC).

- **Concordance – ‘Results’**

These options allow you to adjust the way the concordance output is displayed.

- **Concordance – Results – ‘Filter rows’**

This filter option lets you filter the concordance output by the rows that contain a particular sequence of letters (apart from the node). For example, searching for *hands* in *Oliver Twist* yields 124 results; when we use the option ‘filter rows’ and search for *pockets*, this is filtered down to 8 results as illustrated in Figure 7.

CLiC 1.6:4a999a

UNIVERSITY OF BIRMINGHAM Arts & Humanities Research Council The University of Nottingham UNITED KINGDOM • CHINA • MALAYSIA

Showing 1 to 8 of 8 entries (filtered from 124 total entries) from 1 book

	Left	Node	Right	In bk.
1	he dim light, like high-shouldered ghosts with their	hands	in their breeches pockets. Coffin-plates, elm-chips,	
2	him; and another was so obliging as to put his	hands	in his pockets, in order that, as he was very	
3	as usual; Master Bates sauntering along with his	hands	in his pockets; and Oliver between them, wondering	
4	d the doctor; 'let me think.' Mr. Losberne thrust his	hands	into his pockets, and took several turns up and do	
5	ent to this result of experience, the doctor put his	hands	into his pockets, and walked up and down the room	
6	ed his superiority in a becoming manner, thrust his	hands	into his pockets, and sauntered towards the door,	
7	cut short by the entry of Master Bates, with his	hands	in his breeches-pockets, and his face twisted into	
8	Bumble fixed his hat on very tight, and putting his	hands	in his pockets, followed his helpmate downstairs. Y	

Load | Merge | Save | Clear | About

Concordance

Search the corpora:
Oliver Twist (Charles Dickens) x

Only in subsets:
All text

Search for terms:
hands

Whole phrase Any word

Results

Filter rows:
pockets

Show metadata columns

Figure 7: Concordance of *hands* in *Oliver Twist* filtered down to *pockets* in the co-text.

Note that the filter, when searching for character sequences does not necessarily search for complete words: for example, filtering a concordance of *head* in *Oliver Twist* for *eat* yields both occurrences of the verb *eat*, and the instance *threatened*, which contains the same sequence of letters (see Figure 8). The filter function is cruder than the KWICGrouper; it can be usefully applied to filter down a large set of results before you do a more fine-grained categorisation. You might want to filter down the results to rows containing similar word forms. For example, filtering for *girl* will also retrieve rows containing *girlish* and *girls*. Moreover, unlike the main concordance search and the KWICGrouper, the filter lets you search for particular types of punctuation (e.g. round brackets used in suspensions).

CLiC 1.6:4a999a

UNIVERSITY OF BIRMINGHAM Arts & Humanities Research Council The University of Nottingham UNITED KINGDOM • CHINA • MALAYSIA

Showing 1 to 29 of 29 entries (filtered from 223 total entries) from 1 book

	Left	Node	Right	In bk.
1	is; and, running after him, bestowed a blow on his	head,	which would inevitably have beaten in any skull bu	
2	eyes. His hat was stuck on the top of his	head	so lightly, that it threatened to fall off every mome	
3	boy's face. There was its living copy. The eyes, the	head,	the mouth; every feature was the same. The expre	
4	say: looking at him with great complacency from	head	to foot, that she really didn't think it would have	
5	my death, or I'll be content to eat my own	head,	sir! This was the handsome offer with which Mr. G	
6	pass which will enable a gentleman to eat his own	head	in the event of his being so disposed, Mr. Grimwig	
7	uestion, a very thick coating of powder. 'I'll eat my	head,	sir,' repeated Mr. Grimwig, striking his stick upon t	
8	this bit of peel upon the staircase, I'll eat my	head,	and his too. 'No, no, he has not had one	
9	that boy returns to this house, sir, I'll eat my	head.	With these words he drew his chair closer to the	
10	clamation: 'A beadle. A parish beadle, or I'll eat my	head."	Pray don't interrupt just now,' said Mr. Brownlow. 'I	

Load | Merge | Save | Clear | About

Concordance

Search the corpora:
Oliver Twist (Charles Dickens) x

Only in subsets:
All text

Search for terms:
head

Whole phrase Any word

Results

Filter rows:
eat

Show metadata columns

KWICGrouper

Search in span:

Figure 8: Filtering for the letter sequence *eat* returns forms of the verb *eat* and other words containing the sequence.

○ Concordance – Results – ‘Show metadata columns’

Show the chapter, paragraph and sentence number for each row. (Used to be “Toggle Metadata” until CLiC 1.5). This illustrates where in the book you are and can be the basis for sorting (see section on sorting below).

● Concordance – Basic sorting


The concordance lines can be sorted by any of the columns in the concordance by clicking on the header, which will then be marked with dark arrows. For example, by clicking on “Left” the lines will be sorted by the first word to the left of the node and by clicking on “Right” by the first word on the right. If you have the metadata columns activated you can also sort by these, for example to sort all entries by chapter. Similarly, if you have created your own


tags (see 'tagging' section below), you can sort for lines with a particular tag. Clicking on the same header a second time will reverse the order of sorting.

Note that you can create a 'sorting sequence' by clicking on various headers while pressing the shift key. For example, you could sort a concordance first by the words on the right and then by book, as illustrated in Figure 9, which shows a concordance of *bank* first by book – so that results from *Bleak House* (BH) come first – and then ordered

CLiC
1.6:69376d5

UNIVERSITY OF
BIRMINGHAM

 Arts & Humanities
Research Council

 The University of
Nottingham
UNITED KINGDOM - CHINA - MALAYSIA

Showing 1 to 50 of 287 entries from 15 books

Left	Node	Right	Book	Ch.	Par.	Sent.	In bk.
1 and I am one.We are not rich in the	bank,	but we have always prospered, and we have qu	BH	67	16	56	
2 erest in addition to its real charms.There was a	bank	here, too, which was a famous one for violets; e	BH	36	17	83	
3 venty-six hundred thousand million of parcels of	bank-	notes!""Will somebody give me a quart pot?" e	BH	33	85	207	
4 enty thousand pound down, ready for you in the	Bank	of England if you'll only name one of 'em,' I	BH	11	47	109	
5 sh-bowls, old chairs and mattresses stuffed with	Bank	of England notes.It possesses itself of the sixp	BH	39	59	247	
6 ating the natives of Borrioboola-Gha, on the left	bank	of the Niger."As Ada said nothing, but looked a	BH	4	28	67	
7 persons in each, either gone or going to the left	bank	of the Niger."I thought of the one family so	BH	23	114	316	
8 us who were neither gone nor going to the left	bank	of the Niger, and wondered how she could be	BH	23	115	317	
9 asks the trooper with a glance before him."The	bank,	the factory, or the house?" the workman wants	BH	63	8	13	
10 husband and wife, and with a handsome sum in	bank	(for the locksmith could afford to give his daug	BR	82	12	45	
11 , he was seen.He headed two attacks upon the	Bank,	helped to break open the Toll- houses on Black	BR	67	16	57	
12 the world's ladder--such as the breaking of the	bank	in which her husband kept his money, or some	BR	7	2	4	
13 , he came his accustomed road upon the river's	bank,	intending to pass through Westminster Hall intc	BR	43	9	20	
14 persons of their frugal habits, a goodly sum in	bank.	Moreover they had Grip in company; and when	BR	47	3	9	
15 n, and both Houses of Parliament, the Mint, the	Bank	of England, and the Judges of the land; when	BR	74	3	8	

Figure 9: Concordance of *bank* in novels by Charles Dickens – first ordered by book, then by the first word on the right.

• Concordance – 'KWICGrouper'

The KWICGrouper is a tool that allows you to quickly group the concordance lines according to patterns that you find as you go through the concordance. For a basic introduction to the KWICGrouper functionality (in the CLiC 1.5 interface) you can watch our KWICGrouper video tutorial from May 2017.⁵

The idea of the KWICGrouper is that you look for patterns as you search for particular words. Any matching lines will be highlighted and moved to the top of the screen. Among the matching lines we further distinguish between the lines based on how many matches they contain. A line with one match is highlighted in light green, lines with two matches are coloured in a darker green, those with three in purple and, finally, those with four in pink. (For lines with more matches than these, the colours with repeat.) The KWICGrouper gives you two options:

1. 'Search in span': Set the span for the KWICGrouper search. By dragging the slider you can adjust the number of words that will be searched to the left and right of the search term. The maximum (and default) span is 5 positions to either side.

⁵ The video is available from our CLiC Dickens blog at <https://blog.bham.ac.uk/clic-dickens/2017/06/22/video-introducing-the-clic-kwicgrouper-function-to-group-concordance-lines/>

2. 'Search for types': Choose one or more words to search for in the span. This is currently limited to single words, but there is no limit on how many words you add.

The total number of matching rows will be displayed at the top; the process is illustrated in Figures 10–15. Figure 10 shows the plain concordance lines as returned when searching for *fire* in Dickens's novels.

The screenshot shows the CLiC 1.6.69376d5 interface. The main table displays concordance lines for the search term 'fire'. The table has columns for 'Left', 'Node', 'Right', 'Book', 'Ch.', 'Par.', 'Sent.', and 'In bk.'. The first 13 rows are visible, showing various contexts where 'fire' appears in Dickens's novels.

	Left	Node	Right	Book	Ch.	Par.	Sent.	In bk.
1	at a keeper's lodge and seeing the light of a	fire	upon the latticed panes, and smoke rising from	BH	2	3	23	
2	to my Lady, who is on a sofa near the	fire,	shading her face with a hand-screen."It would b	BH	2	16	83	
3	among the national bulwarks.It happens that the	fire	is hot where my Lady sits and that the hand-scr	BH	2	26	101	
4	mother and I were sitting at the table before the	fire.	The clock ticked, the fire clicked; not another sc	BH	3	9	41	
5	at the table before the fire.The clock ticked, the	fire	clicked; not another sound had been heard in th	BH	3	9	42	
6	ny boxes, I asked him whether there was a great	fire	anywhere?For the streets were so full of dense	BH	3	112	281	
7	it--and politely put an arm-chair for me by the	fire.	He then called my attention to a little looking-gl	BH	3	118	291	
8	deman gave me as he spoke.He then stirred the	fire	and left me.Everything was so strange--the stra	BH	3	121	300	
9	on I went on, thinking, thinking, thinking; and the	fire	went on, burning, burning, burning; and the can	BH	3	122	303	
10	tleman were standing near a great, loud-roaring	fire.	A screen was interposed between them and it, a	BH	3	125	310	
11	in, and I saw in the young lady, with the	fire	shining upon her, such a beautiful girl!With such	BH	3	126	312	
12	sitting in the window-seat, with the light of the	fire	upon us, talking together as free and happy as	BH	3	128	316	
13	sat, he stood by us, in the light of the	fire,	talking gaily, like a light-hearted boy.He was ver	BH	3	130	321	

The sidebar on the right contains search controls. It includes a 'Concordance' section with a search bar, a dropdown for 'Novels by Charles Dickens', and a 'Search for terms' field containing 'fire'. There are also options for 'Whole phrase' and 'Any word', a 'Filter rows' section, and a 'KWICGrouper' section.

Figure 10: The first concordance lines of *fire* in Dickens's novels.

Figure 11 illustrates the process of choosing types (forms of words) from co-text surrounding *fire* in the concordance in order to group the concordance lines. The dropdown only contains those word forms that actually appear around the node term in the specified search span. Therefore, while *sitiwation* is listed here, it wouldn't be listed if we had searched for another node term or used other books; it only appears once in this set in the following context (example 2):

- (2) I don't take no pride out on it, Sammy,' replied Mr. Weller, poking the fire vehemently, 'it's a horrid **sitiwation**. I'm actiwallly drove out o' house and home by it.The breath was scarcely out o' your poor mother-in-law's body, ven vun old 'ooman sends me a pot o' jam, and another a pot o' jelly, and another brews a blessed large jug o' camomile-tea, vich she brings in vith her own hands.' [*Pickwick Papers*, Chapter LI.]

CLiC 1.6.69376d5

UNIVERSITY OF BIRMINGHAM Arts & Humanities Research Council The University of Nottingham

Showing 1 to 50 of 1,706 entries from 15 books

	Left	Node	Right	Book	Ch.	Par.	Sent.	In bk.
1	at a keeper's lodge and seeing the light of a	fire	upon the latticed panes, and smoke rising from	BH	2	3	23	
2	to my Lady, who is on a sofa near the	fire,	shading her face with a hand-screen."It would b	BH	2	16	83	
3	among the national bulwarks.It happens that the	fire	is hot where my Lady sits and that the hand-scr	BH	2	26	101	
4	mother and I were sitting at the table before the	fire.	The clock ticked, the fire clicked; not another so	BH	3	9	41	
5	at the table before the fire.The clock ticked, the	fire	clicked; not another sound had been heard in th	BH	3	9	42	
6	my boxes, I asked him whether there was a great	fire	anywhere?For the streets were so full of dense	BH	3	112	281	
7	it--and politely put an arm-chair for me by the	fire.	He then called my attention to a little looking-gl	BH	3	118	291	
8	tleman gave me as he spoke.He then stirred the	fire	and left me.Everything was so strange--the stra	BH	3	121	300	
9	en I went on, thinking, thinking, thinking; and the	fire	went on, burning, burning, burning; and the can	BH	3	122	303	
10	tleman were standing near a great, loud-roaring	fire.	A screen was interposed between them and it, a	BH	3	125	310	
11	in, and I saw in the young lady, with the	fire	shining upon her, such a beautiful girl!With such	BH	3	126	312	
12	sitting in the window-seat, with the light of the	fire	upon us, talking together as free and happy as	BH	3	128	316	
13	sat, he stood by us, in the light of the	fire,	talking gaily, like a light-hearted boy.He was ver	BH	3	130	321	
14	to talk about, and we talked about it; and the	fire,	which had left off roaring, winked its red eyes a	BH	3	130	324	
15	and sitting in an arm-chair at a table near the	fire,	was his lordship, whose robe, trimmed with bea	BH	3	132	328	
16	y," said Mr. Kenge, standing with his back to the	fire	and casting his eyes over the dusty hearth-rug	BH	4	3	6	
17	made the room taste strongly of hot tallow (the	fire	had gone out, and there was nothing in the grat	BH	4	28	63	

Load | Merge | Save | Clear | Help

Concordance

Search the corpora:
Novels by Charles Dickens

Only in subsets:
All text

Search for terms:
fire

Whole phrase Any word

Results

Filter rows:
e.g. 'hands'

Show metadata columns

KWICGrouper

Search in span:
L5 L4 L3 L2 L1 R1 R2 R3 R4 R5

Search for types:
sit
sit
situation
sits
sittin
sitting

Figure 11: Selecting types related to sitting from the KWICGrouper to group the concordance lines.

CLiC 1.6.69376d5

UNIVERSITY OF BIRMINGHAM Arts & Humanities Research Council The University of Nottingham

Showing 1 to 50 of 1,706 entries from 15 books, 128 entries with 1 KWIC match

	Left	Node	Right	Book	Ch.	Par.	Sent.	In bk.
1	went into the castle.There, we found, sitting by a	fire,	a very old man in a flannel coat: clean, cheerful	GE	25	39	107	
2	ds.As Mortimer Lightwood sat before the blazing	fire,	conscious of drinking brandy and water then anc	OMF	14	62	172	
3	n it was dark, and he sat before the coffee-room	fire,	awaiting his dinner as he had awaited his breakf	TTC	4	25	66	
4	and my still sitting , musing, over the coffee-room	fire.	At last I rose to go to bed, much to	DC	19	118	285	
5	ood pondering, and she sat looking at the dusty	fire	without moving, for some time.But, the moment	OMF	45	69	125	
6	and hailing it.Therefore, he sat before his dying	fire,	sorrowful to think upon the way by which he hac	LD	13	163	413	
7	got to his lodging, he sat down before the dying	fire,	as he had stood at the window of his old	LD	13	159	400	
8	tella's hearing.But, when we sat by her flickering	fire	at night, she was most weird; for then, keeping	GE	38	22	47	
9	pression of having seen Phil sitting by the frosty	fire	staring at him open-mouthed, and gallops away	BH	56	44	138	
10	on remained unrecognised, and sat staring at her	fire,	unobserved.Attracted by a louder sputtering tha	DS	34	3	8	
11	my basket of keys, stop to sit down by her	fire	and accommodate myself for a little while to her	BH	30	37	91	
12	went back to his room, sat down again before his	fire,	and made up his mind that he was glad he	LD	16	98	261	
13	Mr Dorrit's mind, either when he sat down by his	fire	again, or when he went to bed.All night he	LD	53	87	184	
14	looked in, and saw him sitting lonely by his	fire,	brooding over it.He chanced to look up with a	OMF	63	16	42	
15	the clerk of Mr. Grewgious sit and write by HIS	fire.	A pale, puffy-faced, dark-haired person of thirty,	ED	11	10	31	
16	ty-seven.As Mr. Grewgious sat and wrote by his	fire	that afternoon, so did the clerk of Mr. Grewgiou	ED	11	10	31	
17	he raisins were gone, they sat down by the huge	fire	of blazing logs to a substantial supper, and a mi	PP	28	114	253	

Load | Merge | Save | Clear | Help

Concordance

Search the corpora:
Novels by Charles Dickens

Only in subsets:
All text

Search for terms:
fire

Whole phrase Any word

Results

Filter rows:
e.g. 'hands'

Show metadata columns

KWICGrouper

Search in span:
L5 L4 L3 L2 L1 R1 R2 R3 R4 R5

Search for types:
a-sittin
sat
sit
sits

Tags
Selected rows are part of:
Manage tag columns...

Figure 12: The resulting 'KWICGrouper' concordance lines: the selected types are listed in the search box on the right; and in the case of this example it is suitable to restrict the search span to only the left side of the node.

The KWICGrouper only searches through a number of words to the left and right of the node term, as specified by the search span. Figure 12 shows the resulting concordance lines according to the KWICGrouper settings on the right. Apart from the selected search types the search span has also been restricted to the left side so that clearer patterns of sitting by the fire become visible.

The screenshot shows the CLiC 1.6 web interface. At the top, there are logos for CLiC 1.6, the University of Birmingham, and the University of Nottingham. Below the logos, a message states: "Showing 1 to 50 of 1,706 entries from 15 books, 36 entries with 1 KWIC match". The main table displays concordance results for the search terms 'fire' and 'back'. The table has columns: Left, Node, Right, Book, Ch., Par., Sent., and In bk. The first 14 rows are highlighted in green, indicating a KWIC match. The right sidebar contains search controls: 'Search the corpora:' with a dropdown set to 'Novels by Charles Dickens', 'Only in subsets:' with a dropdown set to 'All text', 'Search for terms:' with a text input containing 'fire', and radio buttons for 'Whole phrase' and 'Any word'. Below these are sections for 'Results', 'Filter rows:' with a text input 'e.g. 'hands'', a checkbox for 'Show metadata columns', 'KWICGrouper' with a 'Search in span:' slider, and 'Search for types:' with a dropdown set to 'back'.

Left	Node	Right	Book	Ch.	Par.	Sent.	In bk.
1 > rose, and to dart back into it when the	fire	fell, were sufficient witnesses of his mak	LD	48	7	41	
2 id Mr Dombey, turning back towards the	fire,	'you know what to do with them, I presu	DS	44	33	63	
3 Mr. Kenge, standing with his back to the	fire	and casting his eyes over the dusty hea	BH	4	3	6	
4 Skimpole, standing with his back to the	fire	after giving up the sketch when it was h	BH	6	118	347	
5 urveydrop, standing with his back to the	fire	and waving his gloves condescendingly.'	BH	14	75	206	
6 y were pictures--and comes back to the	fire	repeating, "Sir Leicester Dedlock, Barone	BH	27	6	17	
7 s on the hearth-rug with his back to the	fire,	watching the disappearance of Mr. Smal	BH	27	60	138	
8 n; and Mrs Pipchin, with her back to the	fire,	stood, reviewing the new-comers, like ar	DS	8	91	222	
9 chair, and standing with his back to the	fire.'	Have you anything there for me?" "I don't	DS	13	21	60	
10 so, as he still stood with his back to the	fire,	drawn up to his full height, and looking	DS	13	32	84	
11 eman was standing with his back to the	fire,	and his hands under his coat-tails, looki	DS	13	73	157	
12 d stationed himself with his back to the	fire,	for the purpose of observing the effect	ED	4	39	103	
13 d on the hearth-rug with his back to the	fire,	and went through his favourite action of	GE	40	48	92	
14 y shoulders squared and my back to the	fire."	You have just come down?" said Mr. Dr	GE	43	12	37	

Figure 13: The first lines of *fire* co-occurring with *back* (i.e. one KWICGrouper match) are highlighted and moved to the top.

Apart from looking for characters sitting by the fire, it might also be of interest to look for characters standing by the fire. We have shown in our previous work (chapter 6 of Mahlberg 2013⁶) that the cluster *with his back to the fire* is prominent in Dickens's and 19th century novels by other writers. Figure 13 shows the first concordance lines of *fire* with *back* on the left.

The output from the KWICGrouper lists at the top of the screen the number of lines that contain any number of matches. In the case of Figures 12 and 13 there are only lines with one match, but no lines with more than one match. So, in Figure 13, the message says "36 entries with 1 KWIC match", this means that 36 lines contain both *fire* and *back*. This function becomes useful when we now look for gendered pronouns. As shown in Figure 14, there are 27 lines in which *fire* co-occurs with both *back* and *his* (mostly of the pattern *with his back to the fire*). On the other hand, as we can see from Figure 15 that Dickens's novels contain only instance of *fire* co-occurring with *back* and *her* (*with her back to the fire*).

⁶ Mahlberg, M. (2013). *Corpus Stylistics and Dickens's Fiction*. London: Routledge.

CLiC 1.6-4aa999a

UNIVERSITY OF BIRMINGHAM Arts & Humanities Research Council The University of Nottingham

Showing 1 to 50 of 1,706 entries from 15 books, 27 entries with 2 KWIC matches, 141 entries with 1 KWIC match

Left	Node	Right	Book	Ch.	Par.	Sent.	In bk.
1 Mr. Kenge, standing with his back to the	fire	and casting his eyes over the dusty hear	BH	4	3	6	
2 Skimpole, standing with his back to the	fire	after giving up the sketch when it was h	BH	6	118	347	
3 urveydrop, standing with his back to the	fire	and waving his gloves condescendingly."	BH	14	75	206	
4 s on the hearth-rug with his back to the	fire,	watching the disappearance of Mr. Smal	BH	27	60	138	
5 chair, and standing with his back to the	fire.'	Have you anything there for me?"I don't	DS	13	21	60	
6 so, as he still stood with his back to the	fire,	drawn up to his full height, and looking	DS	13	32	84	
7 leman was standing with his back to the	fire,	and his hands under his coat-tails, looki	DS	13	73	157	
8 d stationed himself with his back to the	fire,	for the purpose of observing the effect	ED	4	39	103	
9 d on the hearth-rug with his back to the	fire,	and went through his favourite action of	GE	40	48	92	
10 m table, and stand with his back to the	fire,	looking at her; she was such an enhance	HT	7	29	75	
11 his drawing-rooms, with his back to the	fire,	waiting for the arrival of his important g	LD	48	7	37	
12 liberty of standing with his back to the	fire,	unless he was quite alone.In the presenc	LD	48	7	38	
13 i, who was lounging with his back to the	fire,	and both elbows on the chimneypiece.'	NN	19	47	92	
14 face--was standing with his back to the	fire	conversing with a younger man, who sto	NN	32	28	67	

Search the corpora: Novels by Charles Dickens

Only in subsets: All text

Search for terms: fire

Whole phrase Any word

Results

Filter rows: e.g. 'hands'

Show metadata columns

KWICGrouper

Search in span: L5 L4 L3 L2 L1 R1 R2 R3 R4 R5

Search for types: back his

Figure 14: The 27 lines with two matches (here, *back* and *his*) are highlighted in a darker green.

CLiC 1.6-4aa999a

UNIVERSITY OF BIRMINGHAM Arts & Humanities Research Council The University of Nottingham

Showing 1 to 50 of 1,706 entries from 15 books, 1 entries with 2 KWIC matches, 103 entries with 1 KWIC match

Left	Node	Right	Book	Ch.	Par.	Sent.	In bk.
1 in; and Mrs Pipchin, with her back to the	fire,	stood, reviewing the new-comers, like an	DS	8	91	222	
2 ouldered within her like an unwholesome	fire.	All closely imprisoned forces rend and de	HT	29	34	89	
3 ell of burnt pens when her feathers took	fire,	which acted as a restorative and woke h	OMF	5	154	378	
4 e rose, and to dart back into it when the	fire	fell, were sufficient witnesses of his mak	LD	48	7	41	
5 I have often known her sit and watch the	fire	- for an hour at a stretch.'"Ay, ay?Has re	HT	19	34	73	
6 breaker, sat with her eyes fixed upon the	fire,	as if she had been deaf to all that passe	OT	19	46	89	
7 in her knees, and with her eyes upon the	fire,	and the fire shining on her ruined beauty	DS	53	111	264	
8 on in a low voice, with her eyes upon the	fire.'	I, with poor father's grave not even clear	OMF	28	127	267	
9 her work abandoned, her eyes upon the	fire,	and her chin upon her hand, 'do you rec	OMF	38	39	97	
10 nan seemed to turn her face towards the	fire	as if to hide her bruised eye; nobody ga	BH	8	84	276	
11 id Mr Dombey, turning back towards the	fire,	'you know what to do with them, I presu	DS	44	33	63	
12 od girl.You had better take her in to the	fire,'	said the locksmith.'Master sounds unfeel	BR	22	16	42	
13 wn at the tea-board, with her face to the	fire.	When once again, in a tone more terrible	MC	25	130	351	
14 a race-horse, with her head close to the	fire,	and her feet pushing the mustard off the	DC	17	87	223	

Search the corpora: Novels by Charles Dickens

Only in subsets: All text

Search for terms: fire

Whole phrase Any word

Results

Filter rows: e.g. 'hands'

Show metadata columns

KWICGrouper

Search in span: L5 L4 L3 L2 L1 R1 R2 R3 R4 R5

Search for types: back her

Figure 15: Only one line contains both *back* and *her*; it is highlighted and shown above single match lines.

- **Concordance – 'Manage tag columns'**

Once you have identified lines with patterns of interest, you might want to place these into one or more categories. CLiC provides a flexible tagging system for this. Figure 16 illustrates the outcome of what a tagged concordance can look like. The tags are user-customised so you can create tags that are relevant to your project. In this case, occurrences of *dream* in *Oliver Twist* have been tagged according to who is dreaming.

CLiC 1.6.69376d5

UNIVERSITY OF BIRMINGHAM Arts & Humanities Research Council The University of Nottingham UNITED KINGDOM • CHINA • MALAYSIA

Showing 1 to 10 of 10 entries from 1 book

	Left	Node	Right	In bk.	dick	you-one	oliver	rose	others
1	I know the doctor must be right, Oliver, because I	dream	so much of Heaven, and Angels, and kind faces th		✓				
2	sy state, between sleeping and waking, when you	dream	more in five minutes with your eyes half open, and			✓			
3	on what seemed to have been a long and troubled	dream.	Feebly raising himself in the bed, with his head re				✓		
4	pity and compassion had awakened some pleasant	dream	of a love and affection he had never known. Thus				✓		
5	all that is going on about us, and, if we	dream	at such a time, words which are really spoken, or				✓		
6	covert for the same reason. 'It must have been a	dream,	Oliver,' said Harry Maylie. 'Oh no, indeed, sir,' repli				✓		
7	erview, which had more the semblance of a rapid	dream	than an actual occurrence, sank into a chair, and				✓		
8	er, and profligacy and riot have staggered home to	dream;	it was at this still and silent hour, that Fagin					✓	
9	of some old friend flashing on one in a vivid	dream.	I need not tell you he was snared away before				✓		
10	and all his recent life had been but a happy	dream.	But it was pure, earnest, joyful reality. They drove				✓		

Previous 1 Next

Only in subsets:
All text
Search for terms:
dream
Whole phrase Any word
Results
Filter rows:
e.g. 'hands'
Show metadata columns
KWICGrouper
Search in span:
Search for types:
Select Some Options
Tags
Selected rows are part of:
dick X
you-one X
oliver X
rose X
others X
Manage tag columns...

Figure 16: Tagged concordance lines of *dream* in *Oliver Twist*.

In order to tag the lines, click on “manage tag columns” (shown in the bottom right corner of Figure 16) and create your own tag(s) through the “Add new” option (see Figure 17). You can rename a tag by selecting it from the “Tag columns” list and renaming it in the text box. Once you have created your tag(s), you can click “back” to return to the menu. Now you can select the relevant concordance lines by clicking on them and you will see that the sidebar contains the list of your tags. Once your lines are selected you can click the tick next to the relevant tag in order to tag the line (see Figure 18). An extra column will appear for each tag and you can sort on these columns as mentioned in the sorting section above. Selected and tagged rows will be automatically deselected when you click on (i.e. select) a new row.

← Back

Tag columns

dick
you-one
oliver
rose
others

Add new Delete selected

Rename selected tag column

Figure 17: The menu for adding and renaming tags.

Showing 1 to 10 of 10 entries from 1 book

	Left	Node	Right	In bk.	click	oliver	rose	others
1	I know the doctor must be right, Oliver, because I	dream	so much of Heaven, and Angels, and kind faces th		✓			
2	my state, between sleeping and waking, when you	dream	more in five minutes with your eyes half open, and		✓			
3	in what seemed to have been a long and troubled	dream.	Feebly raising himself in the bed, with his head re			✓		
4	pity and compassion had awakened some pleasant	dream	of a love and affection he had never known. Thus			✓		
5	all that is going on about us, and, if we	dream	at such a time, words which are really spoken, or			✓		
6	covert for the same reason. 'It must have been a	dream,	Oliver,' said Harry Maylie. 'Oh no, indeed, sir,' repli			✓		
7	review, which had more the semblance of a rapid	dream	than an actual occurrence, sank into a chair, and			✓		
8	if, and profligacy and riot have staggered home to	dream;	it was at this still and silent hour, that Fagin				✓	
9	of some old friend flashing on one in a vivid	dream.	I need not tell you he was snared away before			✓		
10	and all his recent life had been but a happy	dream.	But it was pure, earnest, joyful reality. They drove			✓		

Previous 1 Next

Only in subsets:
All text
Search for terms:
dream
Whole phrase Any word
Results
Filter rows:
e.g. 'hands'
Show metadata columns
KWICGrouper
Search in span:
Search for types:
Select Some Options
Tags
Selected rows are part of:
click X
you-one X
oliver X
rose X
others X
Manage tag columns...

Figure 18: Select a line (by clicking on it) in order to apply an existing tag; once tagged, the tick will appear green for the selected line.

3.3. 'Subsets'

The subsets option can display the full subset of your choice for the selected corpus; click onto the dropdown to select a relevant subset (short suspensions, long suspensions, quotes, non-quotes). For example, you could display all quotes in Dickens's novels for further analysis. You can then use the filter option to narrow down the lines and further search and group the lines using the KWICGrouper. As in the concordance tab, you can create tags to annotate rows with noteworthy patterns.

- **Subsets – Results**

Like in the concordance tab, this allows you to adjust the way the concordance output ('table') is displayed.

- **Subsets – Results – Filter rows**

This filter option lets you filter the concordance output by the rows that contain a particular sequence of letters, as described in Section 3.2 on the filter function in the concordance tab.

Showing 1 to 2 of 2 entries (filtered from 315 total entries) from 1 book

	Left	Node	Right	In bk.
1	said Mrs. Mann. 'Well, well,'	said the beadle, evidently gratified with the compliment;	'perhaps I may be. Perhaps	
2	Perhaps she does see me,'	whispered Oliver, folding his hands together;	'perhaps she has sat by	

Previous 1 Next

Load | Merge | Save | Clear | About

Concordance

Subsets

Search the corpora:
Oliver Twist (Charles Dickens) X

Show subsets:
Long suspensions

Results
Filter rows:
perhaps

Figure 19: Filtering the co-text of long suspensions for *perhaps* in *Oliver Twist*.

Note that the filter searches through the whole row and therefore also accounts for words in the context, not only in the subset itself. For example, when searching through the subset of long suspensions in *Oliver Twist* and filtering rows for *perhaps* the results originate only from the co-text, as *perhaps* does not occur in long suspensions (see Figure 19).

3.4. Clusters

The output of the cluster tool generates frequency lists of single words and clusters (or 'n-grams', where 'n' stands for the length of the phrase). If we choose a '1-gram' (single word), we retrieve a simple word list. (In *Oliver Twist*, for example, the top 10 words retrieved from this tool are *the, and, to, of, a, he, in, his, that* – all function words, as we would generally expect.) A cluster is a repeated sequence of words. CLiC currently supports 1-grams (single words), 3-grams (like *i don't know*) and so on up to 5-grams (*what do you mean by*).

The screenshot shows the CLiC web interface. At the top, there are logos for CLiC 1.6, the University of Birmingham, and the University of Nottingham. The main content area is titled 'Please select a subset' and shows 'Showing 0 to 0 of 0 entries'. Below this is a table with columns 'Cluster' and 'Frequency'. The table is empty, with 'No data available in table' displayed. On the right side, there is a sidebar with navigation links: 'Load | Merge | Save | Clear | About'. Below these are three tabs: 'Concordance', 'Subsets', and 'Clusters' (which is selected). Under the 'Clusters' tab, there are search filters: 'Search the corpora:' with a dropdown menu showing 'Oliver Twist (Charles Dickens)', 'Only in subsets:' with a dropdown menu showing 'Select an Option', and 'n-gram:' with a dropdown menu showing '1-gram' (selected), '3-gram', '4-gram', and '5-gram'.

Figure 20: Cluster options.

As in the other tabs, you can restrict the search to a particular subset so that, for example, you can create frequency lists for clusters in quotes (or the other subsets). You can save the resulting list as a CSV file (for example for use in a spreadsheet viewer) by clicking the 'save' button at the top. Note that the CLiC 'Cluster' tab will display words and clusters with a minimum frequency of 5.

3.5. Keywords

The keywords tool finds words (and phrases) that are used significantly more often in one corpus compared to another. CLiC incorporates the keyword extraction formula reported by Rayson and Garside (2000)⁷. Apart from comparing single words, CLiC also allows you to compare clusters. Whereas the cluster tab focuses only on one corpus, the Keywords function can compare cluster lists. You have to make selections for the following options (also see Figure 21):

⁷ Rayson, P. and Garside, R. (2000). Comparing corpora using frequency profiling. In proceedings of the workshop on Comparing Corpora, held in conjunction with the 38th annual meeting of the Association for Computational Linguistics (ACL 2000). 1-8 October 2000, Hong Kong, pp. 1-6, retrieved from http://ucrel.lancs.ac.uk/people/paul/publications/rg_acl2000.pdf

- **‘Target corpora’**: Choose the corpus/corpora that you are interested in.
- **‘within subset’**: Specify which subset you are comparing (or simply choose ‘all text’)/
- **‘Reference corpora’**: Choose the reference corpus to compare your target corpus to.
- **‘within subset’**: Specify the subset.
- **‘n-gram’**: Do you want to compare single words (1-grams) or phrases (2-grams up to 5-grams)

Keywords

Target corpora:
Select Some Options

...within subset:
Select an Option

Reference corpora:
Select Some Options

...within subset:
Select an Option

n-gram:
1-gram

p-value cut-off:
0.0001

Results

Filter rows:
e.g. 'hands'

Figure 21: The settings for the keywords tab. Here you need to select two sets of corpora for the keyword comparison – target and reference – and their corresponding subsets.

The results are limited to 3000 rows. Generally there will be fewer results. Only overused (positive) keywords are displayed.

Showing 1 to 4 of 4 entries

	N-gram	Target frequency	Ref frequency	Expected target	Expected ref	LL	P
1	what do you mean by	8	14	0.59	21.41	29.82	p < 0.0001
2	not a bit of it	6	6	0.32	11.68	27.11	p < 0.0001
3	have you got to say	5	5	0.27	9.73	22.59	p < 0.0001
4	what have you got to	5	7	0.32	11.68	20.26	p < 0.0001

Previous 1 Next

Target corpora:
Oliver Twist (Charles Dickens) x

...within subset:
Quotes

Reference corpora:
19th century reference corpus x

...within subset:
Quotes

n-gram:
5-gram

p-value cut-off:
0.0001

Figure 22: Key 5-word clusters in Oliver Twist quotes compared to the 19th Century Reference Corpus.

Note that you have to select a subset for each of the two corpora or you'll see the error message: "Please select a subset". So, for example, when comparing 5-grams in *Oliver Twist* (quotes) against the 19th century novels (quotes), we retrieve the results displayed in Figure 22 (for a p-value of 0.0001).

The frequency threshold of 5 used for the cluster tab is not applied to the keyword tab, so that all frequencies are compared. The keyword output shows the top 3000 results (for most comparisons, you will yield fewer results, though). Moreover, CLiC only generates so-called 'positive keywords': those that are 'overused' in the target corpus than in the reference corpus, but CLiC does not generate 'negative' or 'underused' keywords.

4. Appendix: List of texts available in CLiC

Pre-selected corpora in CLiC

DNov	Novels by Charles Dickens
ChiLit	Children's Literature
19C	19 th century reference corpus
Other	A collection of 'other' texts with more set titles from A-Level and GCSE exam specifications.

Table 1: Titles in the DNov, 19C and Other pre-selected corpora.

Title	Author	Corpus
<i>Emma</i>	Jane Austen	19C
<i>Persuasion</i>	Jane Austen	19C
<i>Pride and Prejudice</i>	Jane Austen	19C
<i>Sense and Sensibility</i>	Jane Austen	Other
<i>Mansfield Park</i>	Jane Austen	Other
<i>Northanger Abbey</i>	Jane Austen	Other
<i>Lady Susan</i>	Jane Austen	Other
<i>Lady Audley's Secret</i>	Mary Elizabeth Braddon	
<i>Agnes Grey</i>	Anne Brontë	
<i>The Tenant of Wildfell Hall</i>	Anne Brontë	Other
<i>Jane Eyre</i>	Charlotte Brontë	
<i>The Professor</i>	Charlotte Brontë	
<i>Wuthering Heights</i>	Emily Brontë	
<i>The Last Days of Pompeii</i>	Edward George Bulwer-Lytton	
<i>The Awakening^a</i>	Kate Chopin	Other
<i>Antonina or, the Fall of Rome</i>	Wilkie Collins	
<i>Armada</i>	Wilkie Collins	
<i>The Woman in White</i>	Wilkie Collins	
<i>The Moonstone</i>	Wilkie Collins	Other
<i>Heart of Darkness^b</i>	Joseph Conrad	Other
<i>A Christmas Carol in Prose; Being a Ghost Story of Christmas</i>	Charles Dickens	Other
<i>A Tale of Two Cities</i>	Charles Dickens	DNov
<i>Barnaby Rudge</i>	Charles Dickens	DNov
<i>Bleak House</i>	Charles Dickens	DNov
<i>David Copperfield</i>	Charles Dickens	DNov
<i>Dombey and Son</i>	Charles Dickens	DNov
<i>Great Expectations</i>	Charles Dickens	DNov
<i>Hard Times</i>	Charles Dickens	DNov
<i>Little Dorrit</i>	Charles Dickens	DNov
<i>Martin Chuzzlewit</i>	Charles Dickens	DNov
<i>Nicholas Nickleby</i>	Charles Dickens	DNov
<i>Oliver Twist</i>	Charles Dickens	DNov
<i>Our Mutual Friend</i>	Charles Dickens	DNov
<i>Pickwick Papers</i>	Charles Dickens	DNov
<i>The Mystery of Edwin Drood</i>	Charles Dickens	DNov

<i>The Old Curiosity Shop</i>	Charles Dickens	DNov
<i>Sybil, or the two nations</i>	Benjamin Disraeli	19C
<i>Vivian Grey</i>	Benjamin Disraeli	19C
<i>The Hound of the Baskervilles</i>	Arthur Conan Doyle	19C
<i>The Sign of the Four</i>	Arthur Conan Doyle	Other
<i>Daniel Deronda</i>	George Eliot	19C
<i>The Mill on the Floss</i>	George Eliot	19C
<i>Silas Marner</i>	George Eliot	Other
<i>Middlemarch</i>	George Eliot	Other
<i>A Room with a View</i>	E. M. Forster	Other
<i>Cranford</i>	Elizabeth Gaskell	19C
<i>Mary Barton</i>	Elizabeth Gaskell	19C
<i>North and South</i>	Elizabeth Gaskell	19C
<i>The Yellow Wallpaper</i>	Charlotte Perkins Gilman	Other
<i>Jude the Obscure</i>	Thomas Hardy	19C
<i>The Return of the Native</i>	Thomas Hardy	19C
<i>Tess of the D'Urbervilles</i>	Thomas Hardy	19C
<i>The Portrait of a Lady – Volume 1^b</i>	Henry James	Other
<i>The Portrait of a Lady – Volume 2^b</i>	Henry James	Other
<i>What Maisie Knew</i>	Henry James	Other
<i>Women in Love</i>	D. H. Lawrence	Other
<i>Twelve Years a Slave</i>	Solomon Northup	Other
<i>Frankenstein</i>	Mary Shelley	19C
<i>The Strange Case of Dr Jekyll and Mr Hyde</i>	Robert Louis Stevenson	19C
<i>Dracula</i>	Bram Stoker	19C
<i>Vanity Fair</i>	William Makepeace Thackeray	19C
<i>The Small House at Allington</i>	Anthony Trollope	19C
<i>Adventures of Huckleberry Finn</i>	Mark Twain	Other
<i>The War of the Worlds</i>	H. G. Wells	Other
<i>The Return of the Soldier</i>	Rebecca West	Other
<i>The Picture of Dorian Gray</i>	Oscar Wilde	19C

Notes:

^a The Gutenberg source that we used for *The Awakening* (see <https://github.com/birmingham-ccr/corpora/blob/master/README.pdf>) also contained short stories by Chopin. These have been removed before the main text was imported into CLiC.

^b These titles will be added in mid-November.

Table 2: Titles in ChiLit; a corpus compiled by Anna Cermakova for the GLARE project⁸.

Title	Author
<i>Vice Versa; or, A Lesson to Fathers</i>	F. Anstey
<i>The Brass Bottle</i>	F. Anstey
<i>The Coral Island: A Tale of the Pacific Ocean</i>	R. M. Ballantyne
<i>Peter Pan</i>	J. M. Barrie
<i>The Secret Garden</i>	Frances Hodgson Burnett
<i>Alice's Adventures in Wonderland</i>	Lewis Carroll
<i>Through the Looking-Glass</i>	Lewis Carroll
<i>The Surprising Adventures of Sir Toady Lion with Those of General Napoleon Smith</i>	S.R. Crockett

⁸ See <http://birmingham.ac.uk/glare> for more information on the GLARE project.

<i>Mrs. Overthway's Remembrances</i>	Juliana Horatia Ewing
<i>Jackanapes</i>	Juliana Horatia Ewing
<i>Moonfleet</i>	J. Meade Falkner
<i>Eric; Or, Little by Little</i>	F. W. Farrar
<i>Adventures in Wallypug-Land</i>	G. E. Farrow
<i>The Golden Age</i>	Kenneth Grahame
<i>Dream Days</i>	Kenneth Grahame
<i>The Wind in the Willows</i>	Kenneth Grahame
<i>King Solomon's Mines</i>	H. Rider Haggard
<i>Allan Quatermain</i>	H. Rider Haggard
<i>Winning His Spurs. A Tale of the Crusades</i>	G. A. Henty
<i>With Clive in India; Or, The Beginnings of an Empire</i>	G. A. Henty
<i>Tom Brown's Schooldays</i>	Thomas Hughes
<i>Mopsa the Fairy</i>	Jean Ingelow
<i>Wood Magic, a Fable</i>	Richard Jefferies
<i>Madam How and Lady Why; Or, First Lessons in Earth Lore for Children</i>	Charles Kingsley
<i>The Water-Babies</i>	Charles Kingsley
<i>Stalky and Co</i>	Rudyard Kipling
<i>The Jungle Book</i>	Rudyard Kipling
<i>Prince Prigio</i>	Andrew Lang
<i>At the Back of the North Wind</i>	George McDonald
<i>The Princess and the Goblin</i>	George McDonald
<i>The Three Mulla-mulgars</i>	Walter de la Mare
<i>Masterman Ready; Or, The Wreck of the Pacific</i>	Frederick Marryat
<i>The Settlers in Canada</i>	Frederick Marryat
<i>The Children of the New Forest</i>	Frederick Marryat
<i>Feats on the Fiord</i>	Harriet Martineau
<i>The Crofton Boys</i>	Harriet Martineau
<i>The Peasant and the Prince</i>	Harriet Martineau
<i>The Settlers at Home</i>	Harriet Martineau
<i>A World of Girls: The Story of a School</i>	L. T. Meade
<i>The Cuckoo Clock</i>	Mary Louisa Molesworth
<i>The Tapestry Room: A Child's Romance</i>	Mary Louisa Molesworth
<i>The Carved Lions</i>	Mary Louisa Molesworth
<i>The Book of Dragons</i>	E. Nesbit
<i>The Story of the Treasure Seekers</i>	E. Nesbit
<i>Nine Unlikely Tales</i>	E. Nesbit
<i>The Railway Children</i>	E. Nesbit
<i>Five Children and It</i>	E. Nesbit
<i>The Story of the Amulet</i>	E. Nesbit
<i>The Tale of Peter Rabbit</i>	Beatrix Potter
<i>The Tale of Squirrel Nutkin</i>	Beatrix Potter
<i>The Tale of Benjamin Bunny</i>	Beatrix Potter
<i>The Tale of Two Bad Mice</i>	Beatrix Potter
<i>The Tale of Jemima Puddle-Duck</i>	Beatrix Potter
<i>The Tale of the Flopsy Bunnies</i>	Beatrix Potter
<i>The Fifth Form at Saint Dominic's: A School Story</i>	Talbot Baines Reed
<i>The King of the Golden River; or, the Black Brothers: A Legend of Stiria</i>	John Ruskin
<i>Black Beauty</i>	Anna Sewell
<i>Holiday House: A Series of Tales</i>	Catherine Sinclair
<i>Kidnapped</i>	Robert Louis Stevenson
<i>Treasure Island</i>	Robert Louis Stevenson

Jessica's First Prayer
Little Meg's children
Alone in London
The Rival Crusoes; or The Shipwreck
The Rose and the Ring
Leila at home
The Happy Prince, and Other Tales
The Heir of Redclyffe
The Daisy Chain, or Aspirations
The Dove in the Eagle's Nest
The Little Duke: Richard the Fearless

Hesba Stretton
Hesba Stretton
Hesba Stretton
Agnes Strickland
William Makepeace Thackeray
Ann Fraser Tytler
Oscar Wilde
Charlotte M. Yonge
Charlotte M. Yonge
Charlotte M. Yonge
Charlotte M. Yonge