



Friedrich-Alexander-Universität  
Philosophische Fakultät und  
Fachbereich Theologie



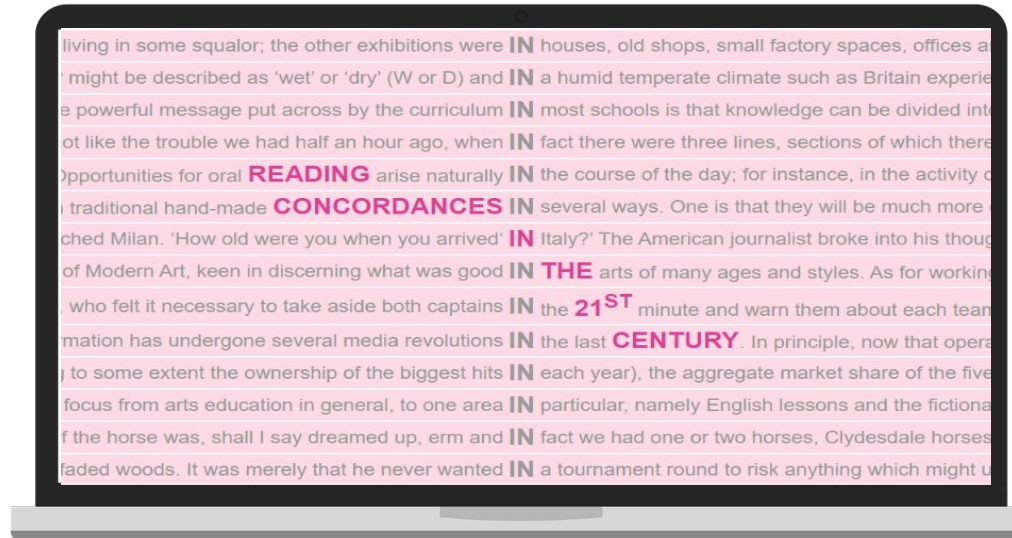
UNIVERSITY OF  
BIRMINGHAM



Deutsche  
Forschungsgemeinschaft



Arts and  
Humanities  
Research Council



3

# Reading Concordances:

a training course in key corpus linguistics methodology

## EESSLI 2025 | 4 – 8 Aug 2025

@schtepf.bsky.social & @michamahlberg.bsky.social

**Reading Concordances in the 21<sup>st</sup> Century (RC21) project team**  
Nathan Dykes • **Stephanie Evert** • **Michaela Mahlberg** • Alexander Piperski

# What we have learned so far from using CQPweb?

- ☐ What strategies have you applied?
- ☐ How do you describe what you do?

# Practices in different disciplines: strategies

- 10 articles / case studies per discipline (40 total)
  - Lexicography
  - Data-driven learning
  - Corpus-assisted discourse studies
  - Literary stylistics
  - post-2000 publication; concrete examples
  - not included: studies using parallel or learner corpora
- Content analysis of methods sections (not literature review)
- Challenge: no systematic standards for describing methodology, mix of tool and method

# Selecting

## Sampling

A random **sample** of 100 concordances of each pair was, subsequently, **closely examined** for semantic prosody. **(Lexicography)**

## Filtering

A concordance **query** for *carry out* with *research* in the context of three tokens to the left and to the right returned 160 hits, and most lines indeed exhibited the collocation *carry out research*, as exemplified in Figure 8, .... **(Lexicography)**

# Selecting

## Cleaning

As we were interested in lexical compounds linked to carbon dioxide rather than carbon as an element, such lexical combinations as, for example, *carbon molecule* or *carbon fibre* were removed. Nonsensical combinations as well as duplicates arising from differences in spelling such as *carbon offset* and *carbon-offset* were also eliminated. **(CADS)**

## Threshold

We analysed the first 500 concordances of *seguire* chosen as “good examples” by the Sketch Engine (cf. Kilgarriff et al., 2008) **(Lexicography)**

# Ordering

## Sorting

WordSmith Tools was mainly used for **sorting** and categorising concordances because it allows the analyst to assign values to lines and to move easily from concordance to text. [Figure 3 shows sorting by 1R] **(CADS)**





The students were then instructed to search for the preceding underlined word *friendship* using IntelliText, and to examine the concordance lines to identify which preposition should follow *friendship* in this context, by **sorting** right or left contexts in keyword-in-context (KWIC) lines. **(DDL)**

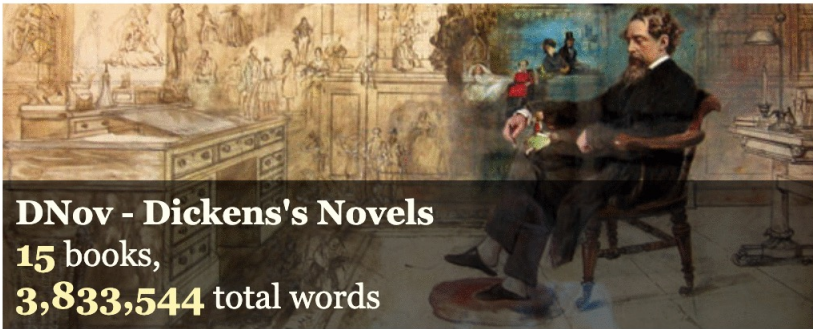
# Ordering

## Ranking

We analysed the **first 500** concordances of *seguire* **chosen** as “good examples” by the Sketch Engine (cf. Kilgarriff et al., 2008)  
**(Lexicography)**

# Corpus Linguistics in Context



**DNov - Dickens's Novels**  
**15** books,  
**3,833,544** total words

**W**elcome to CLiC. The CLiC web app has been developed as part of the [CLiC Dickens project](#), which demonstrates through corpus stylistics how computer-assisted methods can be used to study literary texts and lead to new insights into how readers perceive fictional characters.

The African American Writers 1892-1912 (AAW) Corpus is currently available as a beta release. We welcome any feedback on this corpus, in particular, but also on the CLiC web app as a whole. You can contact us via [email](#).

For more information on how to use CLiC, please read [the CLiC user-guide](#). For the latest information on the CLiC project, look at the [CLiC blog](#) or the [CLiC twitter feed](#). Please also check the [maintenance schedule](#) for dates on which the website will be inaccessible.

Please choose a function in the control bar to the right (click the icon in the top right if it is not displayed).

**Citing CLiC**

Mahlberg, M., Stockwell, P., Wiegand, V. and Lentin, J.  
2020. *CLiC 2.1. Corpus Linguistics in Context*.  
<https://clik.bham.ac.uk>

[clik.bham.ac.uk](https://clik.bham.ac.uk)

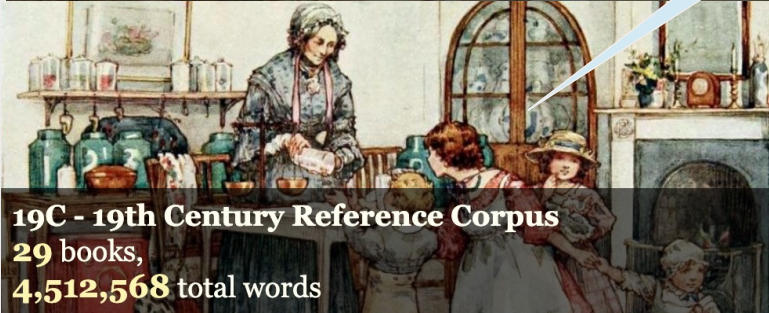
154 texts

16.7 million words



# The texts in CLiC

Such as *Jekyll & Hyde*,  
*Frankenstein*, *Dracula*, *Jane Eyre* and *Pride and Prejudice*



**154 texts 16.7  
million words**



Such as *A Christmas Carol*, *The Yellow Wallpaper*, *The Sign of the Four*, but also some nonfiction like Dickens's *American Notes and Pictures from Italy*

# Reading a story

Once upon a time--of all the good days in the year, on Christmas Eve--old Scrooge sat busy in his counting-house. It was cold, bleak, biting weather: foggy withal: and he could hear the people in the court outside, go wheezing up and down, beating their hands upon their breasts, and stamping their feet upon the pavement stones to warm them. The city clocks had only just gone three, but it was quite dark already--it had not been light all day--and candles were flaring in the windows of the neighbouring offices, like ruddy smears upon the palpable brown air. The fog came pouring in at every chink and keyhole, and was so dense without, that although the court was of the narrowest, the houses opposite were mere phantoms. To see the dingy cloud come drooping down, obscuring everything, one might have thought that Nature lived hard by, and was brewing on a large scale.

*(A Christmas Carol, Stave 1, paragraph 10)*



# Finding a search word?

Once upon a time--of all the good days in the year, on Christmas Eve--old Scrooge sat busy in his counting-house. It was cold, bleak, biting weather: foggy withal: and he could hear the people in the court outside, go wheezing up and down, beating their hands upon their breasts, and stamping their feet upon the pavement stones to warm them. The city clocks had only just gone three, but it was quite dark already--it had not been light all day--and candles were flaring in the windows of the neighbouring offices, like ruddy smears upon the palpable brown air. The fog came pouring in at every chink and keyhole, and was so dense without, that although the court was of the narrowest, the houses opposite were mere phantoms. To see the dingy cloud come drooping down, obscuring everything, one might have thought that Nature lived hard by, and was brewing on a large scale.

(*A Christmas Carol*, Stave 1, paragraph 10)





### Exercise 1.b: The use of "fog\*" across *A Christmas Carol*

1. Open the **CLiC website** in the Chrome browser ([clic.bham.ac.uk](http://clic.bham.ac.uk))
2. Click on the "**Concordance tab**"
3. Under "**Search the corpora**", select "**A Christmas Carol: A Ghost Story of Christmas (Charles Dickens)**" – a fast option to select it is to start typing "Carol" and the title will be recommended.
4. Under "**Search for terms**", enter "**fog\***" (the \* is a "wildcard", a placeholder for extra letters, so you will find words like *foggy*).
5. For each concordance line, click on the slider to see the "**in bk.**" ("**in book**") view. Check the context for each occurrence: how is "fog" used here?
6. Describe the **development** that the different occurrences indicate from the beginning to the end of the novella.

Showing 1 to 9 of 9 entries, Ref. Freq. 315.41 pm, from 1 book

Left	Node	Right	Book	In bk. 4.
1 use. It was cold, bleak, biting weather:	foggy	withal: and he could hear the people in	carol	
2 ears upon the palpable brown air. The	fog	came pouring in at every chink and key	carol	
3 eated himself with rapid walking in the	fog	and frost, this nephew of Scrooge's, th	carol	
4 n was usual with him. ¶ Meanwhile the	fog	and darkness thickened so, that people	carol	
5 the baby sallied out to buy the beef. ¶	Foggier	yet, and colder. Piercing, searching, bit	carol	
6 ed in terror, leaving the keyhole to the	fog	and even more congenial frost. ¶ At ler	carol	
7 was fain to grope with his hands. The	fog	and frost so hung about the black old	carol	
8 old make out was, that it was still very	foggy	and extremely cold, and that there was	carol	
9 he opened it, and put out his head. No	fog,	no mist; clear, bright, jovial, stirring, cc	carol	

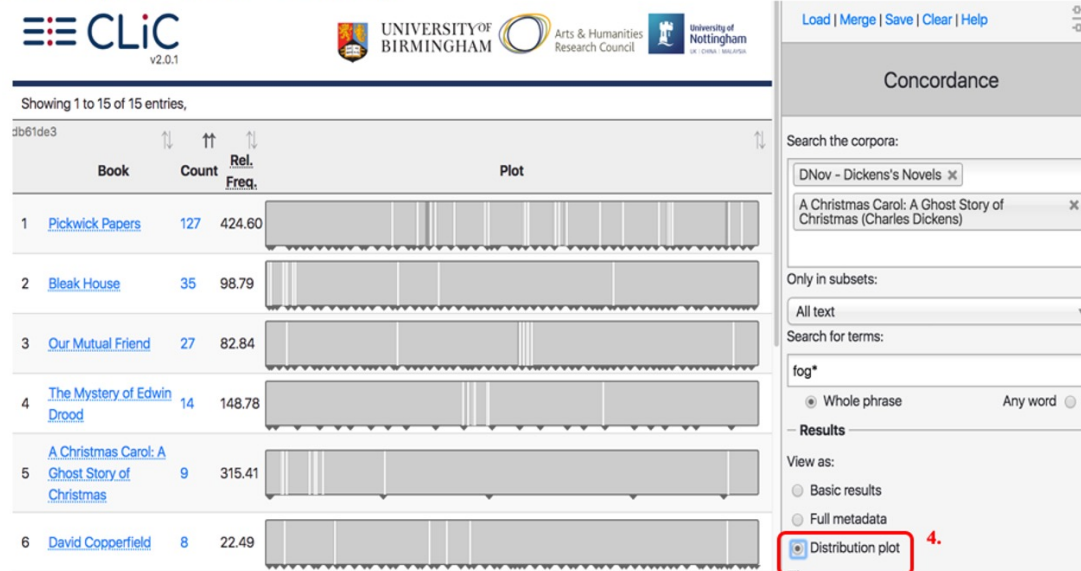
A **concordance** displays all occurrences of a word, phrase, or pattern in a text or set of texts.

Welcome to CLiC. The CLiC web app has been developed as part of the [CLiC Dickens project](#), which demonstrates through corpus stylistics how computer-assisted methods can be used to study literary texts and lead to new insights into how readers perceive fictional characters.

<https://blog.bham.ac.uk/clic-dickens/2019/05/17/finding-links-between-the-structure-and-themes-of-a-christmas-carol/>

### Exercise 1.c: The use of “fog\*” across Dickens's novels & *A Christmas Carol*

1. Stay in the “**Concordance tab**”
2. Under “**Search the corpora**”, keep “**A Christmas Carol**” and add “**DNov – Dickens's Novels**”
3. Under “**Search for terms**”, keep “**fog\***”
4. Under “**Results**” – “**View as**” choose the “**Distribution plot**”. You should then see a plot like this (but with all 15 entries).



A **distribution plot** displays an overview of the occurrences of a word, phrase or pattern across a text or set of texts.

5. How is “fog\*” **distributed** across the texts?
6. In which texts does “fog\*” occur **a lot at the beginning**?
7. **Click on a line** on the distribution plot to **see the context of a particular occurrence**.



## Good Neighbours, Good Friends? Navigating Neighbourhoods, Communities and Connection in Dickens

In this post, Dr Emily Bell (Loughborough University) has used...



CHARLES DICKENS' CAT EXTINGUISHING THE CANDLE.

## "Mew says the cat...Bow-wow-wow says the dog": Which animal did Dickens prefer?

In this post for the "BMI lockdown life" series, Lydia Craig (@lydiaecraig on Twitter) of the Loyola University Chicago delights us with more insights about Charles Dickens, the 16th president of the BMI. Lydia is co-organizer of the upcoming #Dickens150 virtual conference on 9 June. If you want to join this day of virtual talks ... [Continue reading](#)

25 MAY 2020 BY LYDIA CRAIG

1



## Dickens makes the impossible possible: Charles Reporter?

In this guest post, Miriam Helmers (Loughborough University) shows how different digital tools and source between Dickens's journals...

1



## Distance-reading the feminine landscapes of The Awakening

In this guest post, Heather Froehlich, Digital Scholarship Fellow in Text Analysis and Assistant Librarian at Pennsylvania State University, shows that digital humanities is not all about big data but can also provide useful insights on a smaller scale. She demonstrates this with a case study of Kate Chopin's novella The Awakening. One thing we digital ... [Continue reading](#)

2



## Liminality in David Copperfield

In this guest post, Sophie Phelps (University of London) explores how David Copperfield's "childhood" is a study of David Copperfield's "childhood" of characterisation in Dickens...

0



## Dickens, Wooden Legs and the Dickensian Cyberspace

Emma Curry (@EmmalCurry on Twitter) completed her "Language and the Fragmented Body in the Novels of Charles Dickens, nineteenth-century fiction, digital humanities, the history of emotions. During her time at Birkbeck, University of London, she spearheaded the 'Our Mutual Friend Tweets' project, ... [Continue reading](#)

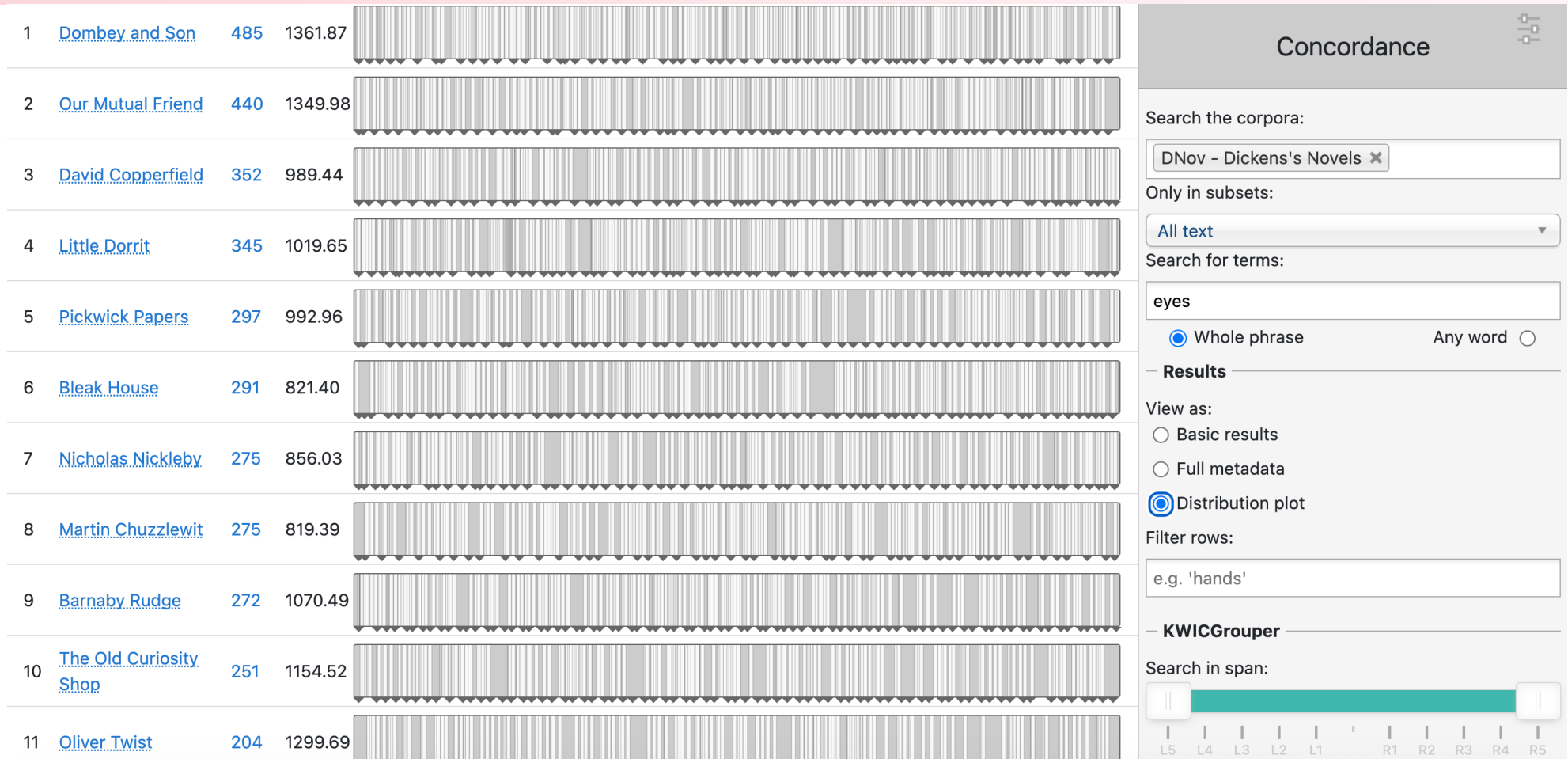
12 APRIL 2018 BY EMMACURRY

at river was like traveling back to the city stream, a great silence, an impenetrable, long, heavy, sluggish silence. There was no joy in the over-shadowed distances. On silvery sand-banks, and buried all day long against the silvery islands; you lost your way on the river, till you thought yourself bewitched. There were moments when one's past came as a shape of an unresisting and noisy dream, and silence and this stillness of this stream a peace.

## "Hello darkness my old friend..." Enacting silence through punctuation in Heart of Darkness

As suggested by the famous song by Simon and Garfunkel, darkness and silence often go together in our imagination. Conrad's novel Heart of Darkness seems to be no exception. The to-and-fro movement from linguistic description to literary appreciation and interpretation is what traditional stylistics and modern computer stylistics have in common, along with the assumption ... [Continue reading](#)

# Eyes in DNov – check the distribution





# Narrow down the context

## Concordance

Search the corpora:

DNov - Dickens's Novels ✕

Only in subsets:

Long suspensions ▼

Search for terms:

eyes



# Discourse levels

## — Highlight subsets —

- ☐ Sentences
- ☒ Quotes
- ☐ Short suspensions
- ☒ Long suspensions
- ☐ Embedded quotes

'I believe nothing is so good for making children lively and cheerful, Sir, as seeing other children playing about 'em,' observed Polly, taking courage.

'I think I mentioned to you, Richards, when you came here,' said Mr Dombey, with a frown, 'that I wished you to see as little of your family as possible.'

'Oh dear yes, Sir, I wasn't so much as thinking of that.'

'I am glad of it,' said Mr Dombey hastily. 'You can continue your walk if you please.'

With that, he disappeared into his inner room; and Polly had the satisfaction of feeling that he had thoroughly misunderstood her object, and that she had fallen into disgrace without the least advancement of her purpose.

Next night, she found him walking about the conservatory when she came down. As she stopped at the door, checked by this unusual sight, and uncertain whether to advance or retreat, he called her in. His mind was too much set on Dombey and Son, it soon appeared, to admit of his having forgotten her suggestion.

'If you really think that sort of society is good for the child,' he said sharply, as if there had been no interval since she proposed it, 'where's Miss Florence?'

Egbert, J. & Mahlberg, M. (2020). Fiction – one register or two? Speech and narration in novels. *Register Studies*, 2(1), 72-101.

Chou, I. & Liu, K. (2023). Style in speech and narration of two English translations of Hongloumeng A corpus-based multidimensional study, *Target*, online.

# Use different functionalities to organise display

Left

↑↓   ↑↓  
Node

Right

⇓  
E

— KWICGrouper —

Search in span:

||

||

L5   L4   L3   L2   L1   '   R1   R2   R3   R4   R5


Search for types:


her ×


— Tags —


# Eye language clic.bham.ac.uk

repV ... -ing ... his / her ... eyes




**UNIVERSITY OF BIRMINGHAM**


**Arts & Humanities Research Council**


**University of Nottingham**  
UK | CHINA | MALAYSIA

---

Showing 1 to 50 of 277 entries, Rel. Freq. 3297.78 pm, from 15 books, 2 lines / 2 books with 3 KWIC matches, 22 lines / 11 books with 2 KWIC matches, 220 lines / 15 books with 1 KWIC match

2a3a120

	Left	Node	Right	Book	In bk.
1	the term I have used'--says Mrs Lammle, fixing her	eyes	on his face, to enforce what she says herself-- 'co	OME	
2	'asked the girl. ¶ 'There ain't,' said Sikes, fixing his	eyes	upon her, and muttering the words to himself; 'ther	OT	
3	¶ "Cousin Richard," said Ada then, raising her blue	eyes	tenderly to his face, "after what our cousin John he	BH	
4	ed, holding the glass above his head, and fixing his	eyes	on Mr Haredale's face; 'a toast to this house and	BR	
5	are a specious fellow,' returned Sir John, fixing his	eyes	upon him, 'and carry two faces under your hood, a	BR	
6	lge of that, mama,' said Annie without removing her	eyes	from his face, 'and he will hear me. If I	DC	
7	u give Mrs Dombey to understand' --the light in his	eyes	fell upon her as he separated his words each from	DS	
8	ing for some time silently rubbing his chin, with his	eyes	fixed upon her, 'and a brave lad, and a good	DS	
9	Alice, taking the hand again, and fixing her lustrous	eyes	lovingly upon her visitor, while she merely addresse	DS	
10	coming to it.' ¶ 'I foresee,' she returned, fixing her	eyes	upon him, 'what it is. But the Lord forbid that	LD	
11	and to--and to see,' said Little Dorrit, raising her	eyes	to his, 'how deeply you mean it, that I cannot	LD	
12	own. ¶ 'Of course,' replied Lord Frederick, fixing his	eyes	on Miss Nickleby, 'how can you a-ask me?' ¶ 'Well,	NN	

DNov - Dickens's Novels x

Only in subsets:  
Long suspensions

Search for terms:  
eyes

☒ Whole phrase ☐ Any word

**Results**

View as:  
☒ Basic results  
☐ Full metadata  
☐ Distribution plot

Filter rows:  
e.g. 'hands'

**KWICGrouper**

Search in span:  

||

||

L5 L4 L3 L2 L1
R1 R2 R3 R4 R5

Search for types:  

fixing x

her x

his x

**Tags**

Mahlberg, M., Wiegand, V., & Hennessey, A. (2020). Eye language – body part collocations and textual contexts in the nineteenth-century novel. In L. Fesenmeier & I. Novakova (Eds.), *Phraséologie et stylistique de la langue littéraire / Phraseology and Stylistics of Literary Language. Approches interdisciplinaires / Interdisciplinary Approaches* (pp. 143–176). Peter Lang.

# Reading Concordances R21 Project

What do users do to organise concordance lines for interpretation? What drives the choices they make, beyond their intuition and the features of concordancing tools they are familiar with?

## Our goals:

- 1. Develop a consistent methodology** of concordance reading strategies and suggesting a transparent terminology that describes tool-independent principles of reading concordances.
- 2. Develop a concordance software framework and new algorithms** that support a consistent concordance reading methodology: the **FlexiConc** library
- 3. Conduct case studies** to illustrate the usefulness of the approach



Released: Aug 4, 2025

6 Aug 2025 | © RC21 Team Reading concordances

# Hands-on:

<https://www.dhss.phil.fau.eu/research/current-projects/reading-concordances-in-the-21st-century-rc21/rc21-events/essli-2025-reading-concordances/>

<https://clic-fiction.com/>

FlexiConc is a flexible Python library for concordance analysis. It enables researchers to study concordances through customisable algorithms and to make their analysis reproducible.

Page 1 of 7

## FlexiConc in CLiC: Basics

This handout illustrates some basic steps of using FlexiConc in CLiC. The examples we chose for illustration deal with patterns of body part nouns in nineteenth century fiction. Our focus is on the functionalities of FlexiConc. We'll provide further references that can support you with the textual analysis.

**FlexiConc** is a flexible Python library for concordance analysis. It enables researchers to study concordances through customisable algorithms and to make their analysis reproducible.

Key features of FlexiConc include:

- **Flexible algorithms** for sorting, ranking, partitioning and clustering concordance lines
- **Modular** and **adjustable** analysis workflows
- **Analysis tree** that tracks all algorithms
- **Integration** with existing corpus tools is possible

How to get FlexiConc

- **PyPI Package:** <https://pypi.org/project/FlexiConc/>

# Fancy contributing to our RC21 blog?

< Research

Current Projects

Reading Concordances in the 21st Century (RC21)

RC21 Blogs


RC21 Events

Publications

Talks

Doctorates

**Concordance Reading × Association Measures**




Concordance Reading × Association Measures Methods to Organize a Dataset of German Support Verb C Reload this page lthor: Xinyao Lu, Friedrich-Alexander-Unviersität Erlangen-Nürnberg Published: 11 June 2025 Introduction How can we collect a dataset of German Support Verb Constructions (SVCs) fr...

11. June 2025 Category: [RC21](#)

Continue >

**Meeting corpus users' needs**




Meeting corpus users' needs Author: Yukio Tono (Tokyo University of Foreign Studies) Published: 17 January 2025 As language tools evolve in the digital age, having the right tools can make all the difference in how we interact with and analyze language. But what makes a tool truly effective...

17. January 2025 Category: [RC21](#)

Continue >

**Expanding Concordance Lines to Aid Interpretability**



Expanding Concordance Lines to Aid Interpretability Author: Mathew Gillings Published: 05 December 2024 Back in June 2024, I had the pleasure of joining and speaking at the RC21 workshop at the ICAME conference in Vigo, Spain. The workshop was designed to stimulate discussion on concordance a...

15. December 2024 Category: [RC21](#)

Continue >

<https://www.dhss.phil.fau.eu/research/current-projects/reading-concordances-in-the-21st-century-rc21/rc21-blogs/>