

# Open Access Resources and Tools for Language Learning

Dr. Jurgita Vaičenonienė  
National coordinator of CLARIN-LT  
Center of Computational Linguistics  
Vytautas Magnus University





<https://pixabay.com/photos/sea-beach-ocean-tropical-paradise-1284467/>

# Corpora and data driven learning

- Digitized naturally occurring texts, usually very large in size (millions of words).
- Processed automatically by a specific software.
- More reliable than personal intuitions about language (even if you are a native speaker).
- Can show trends invisible to the naked eye.
- Present real usage based evidence about language.
- Basic corpus search functions are not complicated and may easily be integrated in language teaching and learning.

# Discovery learning

- Asking a question
- Finding data
- Doing research to see how language is used
- Finding an answer
- Testing the answer with more data
- Sharing the findings
- Asking another question



# Old school

(Vocabulary Your Test. Book 1. 1991. Aura.)

## 33 Puzzle it out

Here are five words:

CABBAGE BANANA CARROT STRAWBERRY EGG

Read through the following dialogue and try to work out which words the above are used instead of (e.g. if you think the word "egg" is used instead of "the" you write "Egg means the" etc.)

A: Excuse me, do you **cabbage** children?

B: No, I'm afraid I **banana**.

A: Oh, **carrot**'s a pity.

B: Why do you say **carrot**?

A: Because everyone should **cabbage** children.

B: Why?

A: Well, **strawberry**'s only natural.

B: I disagree. I **egg carrot** there are far too many children in the world already and I certainly **banana** want to add to the numbers.

A: **Carrot**'s a strange way to **egg**.

B: Is **strawberry**? I **egg strawberry**'s the only sensible way to look at things.

A: Well, I certainly want to **cabbage** at least four children.

B: I **egg carrot**'s a very selfish attitude to take.

A: I **banana** care! **Carrot**'s what I'm going to do.

B: Well — go ahead — **cabbage** all those children but **banana** try to tell me to **cabbage** any.

A: **Banana** worry, **carrot**'s the last thing I'd **egg** of doing.

B: I'm glad to hear **strawberry**.

CABBAGE means .....	BANANA means .....
CARROT means .....	STRAWBERRY means .....
EGG means .....	

# Language data

Collections of texts - corpora

Digitalized books, newspapers, parliamentary debates, historical letters, etc.

Grammars

Termbanks

Translation memories

Digital dictionaries

Multimedia recordings

Tools used to analyse language

# Research Infrastructures

“<...> are facilities that provide resources and services for research communities to conduct research and foster innovation.

They can be used beyond research e.g. for education or public services and they may be single-sited, distributed, or virtual.

They include

- major scientific equipment or sets of instruments
- collections, archives or scientific data
- computing systems and communication networks
- any other research and innovation infrastructure of a unique nature which is open to external users.”

*(European Research Infrastructures)*



# CLARIN: Infrastructure for Humanities and Social Sciences

## CLARIN - European Research Infrastructure for Language Resources and Technology

CLARIN makes digital language resources available to scholars, researchers, students and citizen-scientists from all disciplines, especially in the humanities and social sciences, through single sign-on access. CLARIN offers long-term solutions and technology services for deploying, connecting, analyzing and sustaining digital language data and tools. CLARIN supports scholars who want to engage in cutting edge data-driven research, contributing to a truly multilingual European Research Area. [Read more...](#)



### Participating Organizations

The operations, services and centres of the CLARIN infrastructure are provided and funded by the national consortia in the countries that have joined CLARIN ERIC or by associated centres.

[Learn more](#)



### Services, Tools and Data

CLARIN provides a wide variety of Services and Data sets

[Learn more](#)

 

CLARIN Newsflash January 2021 is Out!



Read the most recent CLARIN Newsflash: January 2021 [here](#)





# How to access data at CLARIN?

- Go to the CLARIN centre of a particular country. For example:
  - [CLARIN-LT](#)
  - [CLARINO](#)
  - [CLARIN-UK](#)
- Search the [Virtual Language Observatory](#)
- Take a look at [Resource families](#)
- and other [services](#)

# How to work with the data?

- Find a corpus in your language of interest.
- Search for the key words.
- Analyse the results.
- Generalize.
- Apply:
  - exercise development,
  - homework assignments (e.g., exemplify rule X; find out how <...>),
  - discussion questions,
  - examples illustrating theory,
  - <...>

# Students as language researchers

- Which word is more frequent in spoken English: “excuse me” or “pardon”? ([BNClab](#))
- Should I put comma before “but”? ([BNC](#))
- Which preposition should be used ? ([BNC](#))
- What are the synonyms for <...>? ([WordNet](#))
- Should I say ‘do’ or ‘make’ homework? ([SketchEngine](#))
- How does word meaning change in different contexts? Does the word ‘wave/banga’ have negative associations in certain context? ([DLKT](#))
- How is Norway/Lithuania represented in UK newspapers? ([WebCorp](#))



The search for 'pardon' found 766 results (55.93 per million).



Usage



Change



Gender



Age



Social Class



Region



Summary



Resources

[Home Page](#)[Choose Language](#)[Choose Corpora](#)[Choose Type of Search](#)[View Results](#)[Build Your Own](#)

Empty

[Parallel Table](#)[Collocation Table](#)[Affix Main Table](#)[Affix Forms Table](#)[Frequency Comparison Table](#)[Frequency Comparison Chart](#)[Keywords/Colloc](#)

## Concordances for but

titleid	left	match	right
<a href="#">A01</a>	the UK and we have had nothing	but	goodwill and people remembering
<a href="#">A01</a>	solo career found mixed success,	but	his role as Billy Flynn in the
<a href="#">A01</a>	after a golden couple of years,	but	the outfit still has a loyal following
<a href="#">A01</a>	might be extended to six months,	but	instead went back to hospital
<a href="#">A01</a>	fiance, Simon Ruddy next year,	but	could not even hold her son on
<a href="#">A01</a>	of chemotherapy on November 1	but	it was just too much for her.
<a href="#">A01</a>	for her. " All was going fine	but	then she got an infection. "
<a href="#">A02</a>	prohibits further illegal development,	but	the amended legislation will also
<a href="#">A02</a>	much more in favour of travellers	but	the only other thing that has

SEARCH

FREQUENCY

CONTEXT

OVERVIEW

(SHUFFLE)

CLICK FOR MORE CONTEXT

☐ [?]

SHOW DUPLICATES

1	KDW	S_conv	A	B	C	(SP:PS1JD) If you were going to do, if you decided you wanted to do research on gypsies (pause) and, you yourself, a lot of gypsies (pause) er,
2	G03	W_fict_prose	A	B	C	've got a bag of change. Those features, do you have to do research for them?' Sometimes. Look, if you can tell me what
3	EBS	W_pop_lore	A	B	C	to hear that their counterparts in Britain have had to defend their right to do research. The arts world does not expect to resolve its funding problems by a change
4	A5Y	W_ac_polit_law_edu	A	B	C	, I can remember when the word went round that someone was coming to do research on the RUC. Like, you must think for yourself; look at the
5	EC7	W_ac_medicine	A	B	C	of the Child Poverty Action Group at the time when you were commissioned to do research and prepare a pamphlet on 'Families with low income in London.' I
6	CLP	W_ac_soc_science	A	B	C	successful conclusion for both sides depending on their level of skill. Attempting to do research on the interview per se is doomed to failure because the skills of the
7	CMF	W_ac_soc_science	A	B	C	Barker (1984) had the unusual experience of being approached and invited to do research into a very enclosed group in which she was already interested, the Unific
8	B7C	W_non_ac_nat_science	A	B	C	the 1930s. In law and in fact every academic has the right to do research, and the universities provide all faculty members with small-scale financial support for this.
9	FYX	W_non_ac_nat_science	A	B	C	But from the age of thirteen or fourteen, I knew I wanted to do research in physics because it was the most fundamental science. This was despite the fact
10	FYX	W_non_ac_nat_science	A	B	C	they asked me about my future plans. I replied that I wanted to do research. If they gave me a first, I would go to Cambridge. If
11	FYX	W_non_ac_nat_science	A	B	C	I therefore applied to do a Ph.D. with Hoyle. My application to do research at Cambridge was accepted, provided I got a first, but to my annoyance
12	FYX	W_non_ac_nat_science	A	B	C	prepared to offer me a sequence of positions in which I only had to do research without having to lecture. We were also fortunate in housing. When we were
13	B2T	W_non_ac_polit_law_edu	A	B	C	to governing body; employers use academics as consultants and engage academic departments to do research for them; there are sponsored research and collabor
14	HCC	W_non_ac_soc_science	A	B	C	. # Regional Policies on Conducting Research in Education # Do you want to do research in schools -- especially in schools other than your own? Due to the increas,
15	HCC	W_non_ac_soc_science	A	B	C	demands on schools some Regions have been developing procedures which those who want to do research in schools must follow. Whilst the procedures may in so
16	HCD	W_non_ac_soc_science	A	B	C	worthwhile endeavour! # Teacher Research and the Regions # Do you want to do research in schools -- especially in schools other than your own? Due to the increas
17	HCD	W_non_ac_soc_science	A	B	C	demands on schools some Regions have been developing procedures which those who want to do research in schools must follow. Whilst the procedures may in so
18	HJ1	W_non_ac_soc_science	A	B	C	University of Minnesota spent a year in Europe. His original intention was to do research in Political Geography, but Central Europe was at that time too political for :
19	HJ2	W_non_ac_soc_science	A	B	C	or the role of the United Nations. Now they are more likely to do research on narrow topics like arms control. In total, much more is known but
20	ADR	W_biography	A	B	C	goes through a number of changes which is very interesting. I tried to do research on the 1950s but it was quite hard to find general information about the day-to-da
21	GTR	W_biography	A	B	C	for acting and theatrical design. After graduation he was awarded a scholarship to do research for a doctoral thesis on Thomas Gray n.v. apparently never complete

## WordNet Search - 3.1

- [WordNet home page](#) - [Glossary](#) - [Help](#)

Word to search for:

Display Options:

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations

Display options for sense: (gloss) "an example sentence"

### Adjective

- [S:](#) (adj) **happy** (enjoying or showing or marked by joy or pleasure) "*a happy smile*"; "*spent many happy days on the beach*"; "*a happy marriage*"
  - [see also](#)
  - [similar to](#)
  - [attribute](#)
  - [antonym](#)
  - [derivationally related form](#)
- [S:](#) (adj) [felicitous](#), **happy** (marked by good fortune) "*a felicitous life*"; "*a happy outcome*"
  - [similar to](#)
  - [derivationally related form](#)
  - [antonym](#)
- [S:](#) (adj) [glad](#), **happy** (eagerly disposed to act or to be of service) "*glad to help*"
  - [similar to](#)
  - [antonym](#)
- [S:](#) (adj) **happy**, [well-chosen](#) (well expressed and to the point) "*a happy turn of phrase*"; "*a few well-chosen words*"
  - [similar to](#)
  - [antonym](#)





# CONCORDANCE

British Academic Written English Corpus (BAWE)



simple homework 11 (1.32 per million)



☐ Details

Left context

KWIC

Right context

1	<input type="checkbox"/>	<input type="checkbox"/>	Narrative recou... they are struggling until you actually mark their books. </s><s> After teaching them for 4 lessons I set a homework and then took their books in. </s><s> I found what they ha
2	<input type="checkbox"/>	<input type="checkbox"/>	Narrative recou... n mistakes was fascinating aswell as misconceptions that individuals had. </s><s> The first part of the homework was giving approximate answers to expressions like 436 :
3	<input type="checkbox"/>	<input type="checkbox"/>	Narrative recou... k but I felt I could as it was a one off. </s><s> This taught me that especially when they are at this age, homework 's set from a textbook can certainly be misinterpreted how
4	<input type="checkbox"/>	<input type="checkbox"/>	Narrative recou... hemselves clearer and also teachers really have to make it clear exactly what is required of pupils with homework at this age. </s><s> Because I had spent a lot of time with
5	<input type="checkbox"/>	<input type="checkbox"/>	Essay employment of part-time and temporary work, short-term contracts, self-employment, contracting out, homework and franchising. </s><s> Of these, the most popular is the
6	<input type="checkbox"/>	<input type="checkbox"/>	Essay they otherwise deserved (Ferri and Liu 2005). </s><s> Most of the emerging countries have done their homework . </s><s> Even without formal imposition, they have soug
7	<input type="checkbox"/>	<input type="checkbox"/>	Case study e and efficient than the traditional test. </s><s> Also, people can have school at home and submit their homework online. </s> The effect of the Internet on business and go
8	<input type="checkbox"/>	<input type="checkbox"/>	Design specific... peal a new segment of the market, which is teachers that are looking for interactivity exercise to do as homework , and a simple way to look after the pupil progresses than
9	<input type="checkbox"/>	<input type="checkbox"/>	Design specific... <s> The Think History! </s><s> Website will create in pupils the link between the class lessons and the homework , using a clear and simple language. </s><s> This is the f
10	<input type="checkbox"/>	<input type="checkbox"/>	Essay 05). </s><s> Relevance to the present is not often established in American English so Did you do your homework ? </s><s> Or I just came home would be considered corre
11	<input type="checkbox"/>	<input type="checkbox"/>	Narrative recou... ss on speaking, reading and listening while writing (and sometimes reading as well) is often part of the homework . </s><s> Especially writing is very time-consuming, beca

Rows per page: 20

Saugoti rezultatus: ☐ Su šaltiniais ☒ Saugoti

Atverti šaltinių sąrašą naujame skirtoke

Rodyti šaltinių sąrašą

Rasta: 832

### Paieška su kontekstiniu žodžiu | Paslėpti paieškos parametrus

- [1] lt Dėl premijos panaikinimo Vengrijoje – protestų  
[2] i Azija (pvz., H.Šlymano atrasta VI Troja)); kita  
[3] izika moko, jog šviesos kvantas yra ir dalelė, ir  
[4] krusteli, ir per visą kūną iki paširdžių nusirita  
[5] kurias mes neseniai iėjome. Į kambarį su staigia  
[6] pietus; ikandin sekė dar viena, galvijus genanti,  
[7] sibaigus arba darbuotojų prašymu. Šiauliuose UAB „Banga  
[8] arąją parą vandens lygis nekilo. Kai kur potvynio  
[9] ario gėlynai dabar jau pavėsingi ir niūrūs. Žiedų  
[10] kme. Nauja daugiausia komercinio pobūdžio statybų  
[11] valdanti PET gamyba užsiimančią bendrovę "Nemuno  
[12] omentarai ir pokalbiai Kasdien mus užlieja nauja  
[13] vau niekas, tik trumpai upės paviršiuje gyvenanti  
[14] iau, vildamasis, kad kokia nors netikėta srovė ar  
[15] pastebėjo valti, – tariau aš. Šūstelėjusi didžiulė  
[16] etuvai. Tarp klausytojų nuvilnijo nepasitenkinimo  
[17] gyvulių kūnais užliejo grėsminga vandenyne kilusi  
[18] ir pažangą dar ir ta prasme, kad kiekviena nauja "Banga"  
[19] askui lūpomis nulaižė nutiškusių lašus. Palaimos  
[20] viesos banga ritmiškai ristųsi per miestą. Gerai,  
[21] 2, G.Bulašas 1, R.Žalys 1, A.Budginas 1. GARGŽDŲ "BANGA"  
[22] vėl atsiduso pelkė, vėl atsirito iki manęs vėsi banga... Po šio atodūso išgirdau sunkius žingsnius. Ka  
[23] 18.30 val. Tauragės "Tauras" (28 tšk.) - Gargždų "Banga" (38 tšk.; 0:3, 2:1 ir 1:2), 19 val. "Lokomotyvo"

## Results for query "Norway"

case insensitive,  
using the Bing (Cognitive) API,  
from sites: mirror.co.uk thesun.co.uk dailystar.co.uk

### 1) <https://www.thesun.co.uk/news/13758376/covid-norway-vaccine-deaths-pfizer-nursing-home/>

Text, Wordlist, text/html, UTF8 (Content-type), date unknown

```
1:  ALARMED' Covid Pfizer vaccine: What happened in Norway? Danielle Cinone 15 Jan 2021, 21:33 NEARLY two
2:      injected with the Pfizer Covid vaccine in Norway - and at least half of them were older than age
3:  very little risk." 4 At least 23 people died in Norway just days after getting the Pfizer coronavirus
4:  reports.Credit: AFP or licensors What happened in Norway? At least 23 people have died in Norway just
5:      in Norway? At least 23 people have died in Norway just days after getting the Pfizer coronavirus
6:  coronavirus-related deaths have been reported in Norway. 4 'The Norwegian Medicine Agency confirms the
7:      The deaths come as more than 30,000 people in Norway have received the first jab of the Pfizer or
8:      home population. A news release states: "In Norway we are now vaccinating the elderly and people in
9:      well tolerated in all age groups." Spate of Norway Covid vaccine deaths 'expected' - health chief
10:     Topics Coronavirus Explainers US Explainers Norway Comments are subject to our community guidelines,
```

### 2) <https://www.thesun.co.uk/news/13757257/norway-pfizer-covid-elderly-deaths/>

Text, Wordlist, text/html, UTF8 (Content-type), 2021-01-16 (Body near 'Last Modified')

```
11:     Politics Opinion Health News 'NOT ALARMED' Norway adjusts Covid vaccine advice on who gets jab
12:  for the latest news & updates 2 Health chiefs in Norway have now adjusted their vaccine adviceCredit:
13:  some frail patients." More than 30,000 people in Norway have received the first jab of the Pfizer or
14:  coronavirus-related deaths have been reported in Norway. And health officials noted that around 400
```

# Create and visualize your data.

<https://voyant-tools.org/>



# Sources

*European Research Infrastructures.*

[https://ec.europa.eu/info/research-and-innovation/strategy/european-research-infrastructures\\_en](https://ec.europa.eu/info/research-and-innovation/strategy/european-research-infrastructures_en)

Hyland, K. 2006. *English for Academic Purposes*. Routledge, London & New York.

McEnery, T & Xiao, R. 2010. What Corpora Can Offer in Language Teaching and Learning. In E Hinkel (ed.), *Handbook of Research in Second Language Teaching and Learning*. vol. 2, Routledge, London & New York, pp. 364-380.





Thank you!