Introducing Sketch Engine for academic prose

Getting started discussion questions

1. Is the word ¹ whose used with things as well as pe	eople?
--	--------

2. Is the word data used as a singular or plural noun?
Singular: The data is
Plural: The data are
Both - it depends on the regional variety of English. e.g., US, UK
Both - it depends on the context, e.g., popular use, scientific use. ²
3. Research³ is not the object of all of these verbs. Underline the ones that you think are correct.
conduct research
make research
design research
do research
create research
cause research
publish research
determine research
support research

Do these verb collocates of research have different subjects? In other words, are they performed by different types of people or things?

4. Do sentences begin with First or Firstly in academic prose?

Registration

MU students go to ske.fi.muni.cz and click on SSO authentication.

¹ ske.li/bnc_whose_500

² ske.li/bnc_sci_data_500, ske.li/bnc_datum_3ps
The CQL search that yields these findings: [lemma = "datum"][tag = "V.Z"]

³ ske.li/bnc ws research n

Simple Query field

Use the Simple Query field to search for words and phrases. The search will find words in all their forms, e.g. singular and plural, verbs in 3rd person singular, past forms, etc.

Question 1. Does the truth lie?

What forms are these two words in, and what is typically used before and after? Is it used in academic prose?



Question 2. Do we say people who or people that?

After searching for both, click on Text Types under Frequency in the left panel. What do you observe?

Question 3. Which of these phrases are used in academic prose?

Use lemmas (base forms) of words. Asterisks are also useful, e.g.,

to * surprise

to * * surprise.

To make a summary list of findings, click on Node forms in the left panel.

What do you learn about their typical contexts?

to my great surprise, is defined as, can(not) be compared with/to, bear in mind. it revolves around, it turns out that, to the best of my knowledge

Phrases

Use the phrase field to study phrases in their exact wordings.

Question 4. What follows in the beginning and at the beginning?

Type in one of these phrases, while your colleagues types in the other. Click Sort Right in the left panel. What do you notice?

Question 5. What is the difference between *Given that* at the beginning of a sentence, with and without a comma?

The use of capital "G" is important. Right sort the concordance page. Does this have a counterpart in spoken language?

Question 6. Which of these expressions are used in academic prose?

Try your intuition first. Choose several to examine.

last but not least, to some extent, the fact that, all of a sudden, as a matter of fact, at the expense of, by the way, in this day and age, in turn, in a row, in the scheme of things, one way or another, in the context of, works in mysterious ways, within reason, well within, at the risk of, provided that

What are they used to express? Are there any patterns in what precedes and follows them?

Word form

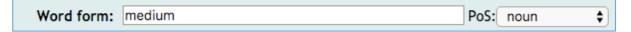
Use the word form field to search for a specific word form, e.g. goings (noun)

Word focus: medium.

Starting with a Simple Query:

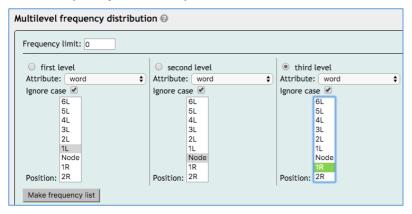
Node tags tell us that it has two parts of speech, mostly _____ as well as _____.

What does **Node forms** tells us about its singular and plural noun forms?



Search for the noun *medium*, in the word form field.

Click Frequency in the left panel and set it like this.



Observations

What do you think the next words might be?

Click on P to find out. Then Right sort the concordance page.

What do you find by Left sorting the page?

Click the column headings to sort them. Set the Frequency limit to 5. What do you observe as you sort the left and the right columns?

What does the high frequency of punctuation suggest?

What have you learnt about language and about using Sketch Engine through studying medium?

Using similar procedures, what can you find out about the words *factor* and *factors* and/or data and *datum*?

Lemma field

Use the lemma field to search for a word and all its word forms. Use lower case only. Specify part of speech.

Question 7. Here are some words that commonly appear in academic prose. What can you learn about how they are used in academic texts?

After searching for a word, you can make a Sample, sort the concordance page left and right, make frequency charts. Some of the features of words we can ask about include:

Are adjectives used attributively, i.e. on the left of the noun e.g. secure income, and/or predicatively, i.e. after a verb usually, i.e. after a verb usually, be, e.g. the results are conclusive.

What prefixes make these words negative? Do the negative adjectives collocate with the same nouns?

Noun	Verb	Adjective	Adverb
problem	support	challenging	potentially
situation	cause	transitional	significantly
attempt	exploit	logical	specifically
strategy	minimise	symbolic	intensely

Lemma:	transitional	PoS:	adjective	•

Character field

The character field can be used to observe punctuation.

Question 8. How much information occurs between brackets?

Search for a left bracket using the character field. What types of things occur between the (and the)?

Question 9. Are colons followed by capital letters?

Search for the colon and observe what follows.

Question 10. How do sentences begin in academic prose?

Enter a full stop, then use Frequency to make this list⁴.

Do they mostly begin with Organisation/Orientation language?

Context

This second level search allows us to limit our top level search items by words and parts of speech.

Question 11. someone ... they - see Q 147.

Enter someone in to Simple Query and they into Context Lemma Filter, allowing 3 tokens to the right. In which sentences does the plural they refer to the singular someone?

James Thomas (2016)

-

⁴ http://ske.li/arc 4gram sentence beginnings

Question 12. Is not only always completed with but also?

Search for *not only* in the Simple Query or Phrase field and observe how many hits there are. Then repeat the search adding *but also* in the Context Lemma Filter field.

Question 13. What's the phrase with ...

```
...term ... understood
... term ... embodies
... difficult to define ...
last ... least
```

Question 14. When cause is used with another verb, is it the to-infinitive form?

In this search, the verb in Simple Query is followed by *to* up to 3 tokens to the right and a verb up to 4.

Simple query: cause Query types Context Text types	Make Concordance
Context Lemma filter	PoS filter
Window: right \$ 3 \$ tokens. Lemma(s): to all \$ of these items.	Window: right \$ 4 \$ tokens. PoS: particle all \$ of these items. preposition pronoun
Make Concordance Clear All	verb ☐ full stop

To search for a verb followed by a verb without to in between, select none of these items.



In a large corpus, there will be very many hits. To make this manageable, make a sample of the findings, say 500. Then Right Sort the concordance page to see the patterns. What patterns do you observe?

The same question can be asked of many common verbs. For example, *sb makes sb do something* or *sb makes sb to do sth*.

Also, let, force, require, encourage, enable, know how, expect, manage. And many more.

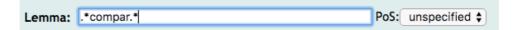
When looking at the results, it is useful to see if there are any typical nouns and adverbs occurring with the verbs. These pairs are known as collocation.

The language of comparison

An academic writing task: compare the abstract we studied in class with one that you've read in your academic work.

The language of **comparison** starts with the word *compare*. Create the **word family** of this word using these operators.

James Thomas (2016)



Then summarise the results using the Word Forms button⁵.

This data can also be restructured using $lempos^6$. Clicking on **P** beside the words in these lists, and then sorting the pages shows the **phrases** that are part of, e.g. Comparisons of X and Y ... (at the start of a sentence), As a comparison, ..., based on a comparison between ...

Word sketch

When determining how the abstract relates to the guidelines for abstract writing, look at the collocations of words such as *advice*, *rules*, *principles*, *guidelines*. For example, using this abbreviated word sketch of advice⁷, you can see how the word combines with other words in various grammatical relationships.

advice (noun) British National Corpus (TreeTagger) freq = 9,689 (85.75 per million)														
object_	subject_o	<u>f</u>		adj_subje	ct_of		modifier	modifies						
	3,755	3.40		828	1.00		<u>154</u>	1.80		3,773	1.30		<u>793</u>	0.30
seek +	<u>365</u>	9.04	regard	<u>15</u>	4.75	fatal	<u>3</u>	5.73	legal +	<u>397</u>	9.49	bureau	31	8.67
offer +	330	8.02	please	<u>5</u>	4.18	available	46	4.88	professional +	<u>155</u>	8.40	agency	54	6.67
heed	<u>30</u>	7.88	influence	3	3.26	sensible	3	4.80	sound	<u>73</u>	8.33	guidance	<u>19</u>	6.65
ignore	<u>63</u>	7.43	suggest	7	3.24	excellent	3	3.74	independent +	140	8.21	worker +	<u>107</u>	6.53
give +	746	7.35	seek	<u>5</u>	3.05	free	<u>7</u>	3.26	practical +	<u>112</u>	8.18	centre	98	6.23
obtain	<u>64</u>	6.67	apply	<u>6</u>	2.90	relevant	<u>3</u>	3.06	medical +	<u>114</u>	8.18	session	28	6.14

Matching activity

object of	what advice does
subject of	types of advice
adj subject of	advice and another noun
modifier	sb or sth verbs advice
modifies	advice is

More grammatical relationships

and/or			pp_on-i			pp_obj_o	<u>f-i</u>		pp_obj_for-	1		pp_of-i		
	1,909	1.70		<u>807</u>	8.90		<u>699</u>	1.50		<u>509</u>	5.00		<u>478</u>	1.00
assistance +	<u>142</u>	8.93	contraception	<u>5</u>	6.87	giver	4	7.22	pharmacist	7	8.09	Clerk	<u>3</u>	6.27
guidance	<u>77</u>	8.43	suitability	4	6.45	plenty	21	6.38	grateful	<u>13</u>	6.51	psychiatrist	<u>3</u>	5.86
help +	153	8.15	eating	3	5.85	piece	67	6.32	Support	<u>5</u>	6.51	counsel	4	5.79
encouragement	25	7.57	diet	14	5.67	spite	7	6.04	ask	80	5.75	therapist	<u>3</u>	5.61
information +	<u>197</u>	6.70	cultivation	<u>3</u>	5.65	source	63	5.94	request	<u>10</u>	5.21	expert	<u>18</u>	5.61
support +	<u>138</u>	6.66	tariff	3	5.18	recipient	4	5.70	eligible	3	5.19	solicitor	11	5.39

For example, such collocations are the core of clauses:

sb is grateful for your/any advice,

the advice of a ___ expert,

sb gives/offers advice and encouragement.

⁵ http://ske.li/bnc_compare_wf

⁶ http://ske.li/bnc_compare_wf_lempos

⁷ http://ske.li/c4u

Now it's time to make and use our own corpora

"How people use words to make meanings" Field-based Article-based AWL nouns words words analogy collocation double helix entity benefit connotation introspection factor corpus category mush hierarchy concept frequency ontology ratio context lemma peacock role element lexicography cognitive profile structure Article by Patrick Hanks

What are the key words and phrases in your corpus and how are they are used?

- · general academic words
- field based words
- article based words
- · discourse markers/signposting
- use of hedging
- first person

http://bit.ly/versatile_deske

