

Sketch Engine 100 Qs



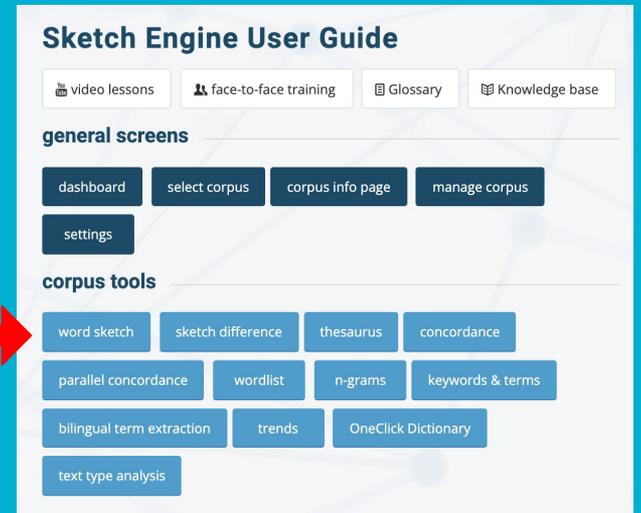
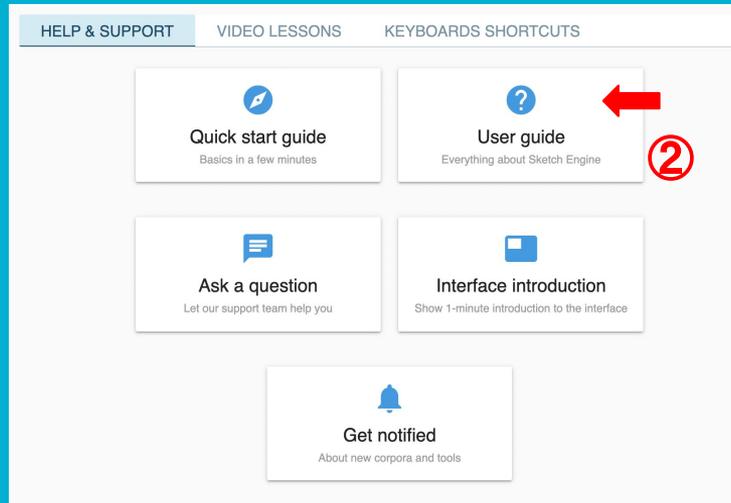
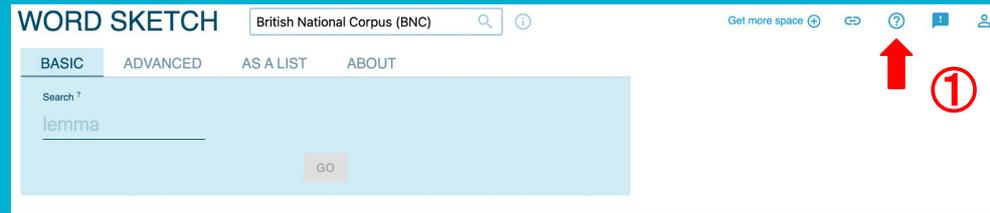
**SKETCH
ENGINE**

初級	Q1 ~Q31
中級	Q32~Q88
上級	Q 89~Q100

初級編

**Q1. Word SketchのUser Guideを概観して
概要を把握しなさい。**

Q1. Word SketchのUser Guideを概観して概要を把握しなさい。



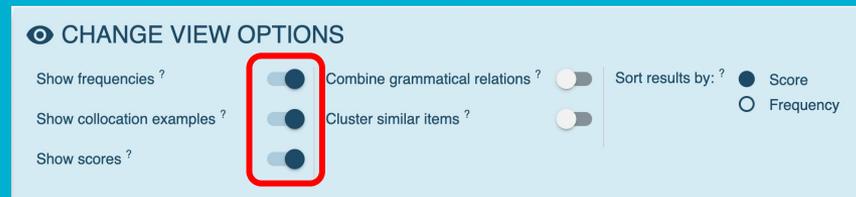
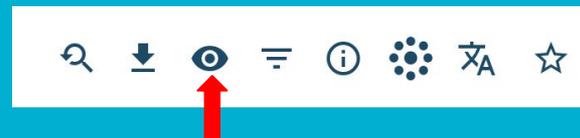
Q2. Word Sketch: BNC利用
Basicにて"think"を検索した後、Change
view options を開いてfrequencies,
cllocation examples, scoresを表示しなさい。

Q2. Basicにて"think"を検索した後、Change view options を開いて frequencies, cllocation examples, scoresを表示しなさい。

Show scoreとは?

The typicality score ([LogDice](#)) indicates how strong the collocation is. The higher the score, the stronger the collocation. A low score means that the words in the collocation also frequently combine with many other words.

共起関係性の強さを示した数値です。数値が高いほど共起関係性が強く、数値が低いほど他の語とも共起することを示しています。



Q3. Word Sketch: BNC利用

Change view optionsを開き “Sort results by” の選択を変更することで検索結果の表示が変わることを確認しなさい。

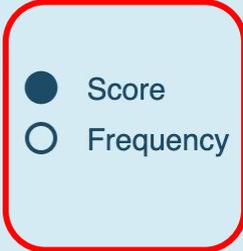
Q3. Change view optionsを開き“Sort results by”の選択を変更することで検索結果の表示が変わることを確認しなさい。

 CHANGE VIEW OPTIONS

Show frequencies ? Combine grammatical relations ? Sort results by: ? Score Frequency

Show collocation examples ? Cluster similar items ?

Show scores ?



Q4. Word Sketch: BNC利用

Advancedにて "research"と入力しPart of speechをnoun、Text typesをWrittenと指定し、"research"後に最も高頻度に使われる前置詞を確認しなさい。

Q4. Advancedにて "research"と入力しPart of speechをnoun、text typeをWrittenと指定し、researchに続く最も高頻度に使われる前置詞を確認しなさい。

WORD SKETCH British National Corpus (BNC)

BASIC **ADVANCED** AS A LIST ABOUT

Search ? **research** Subcorpus ? none (the whole corp...)

Part of speech ?

- auto
- noun**
- verb
- adjective
- adverb
- pronoun

Text types (1) ? ^

Text type ^

- Written books and... x
- Written miscellan... x
- Written-to-be-spo... x

Spoken context-governed

Spoken demographic



GO

prepositional phrases	
... of "research"	11.0%
... for "research"	4.1%
"research" in ...	3.4%
... in "research"	3.0%
"research" into ...	2.7%
"research" on ...	2.5%
... to "research"	1.5%
... on "research"	1.5%
... from "research"	0.8%
... by "research"	0.7%
"research" by ...	0.6%
"research" at ...	0.5%

Q5. Word Sketch: BNC利用

Advancedにて "research" と入力しPart of speechをverb、Text typesをSpokenと指定し、"research"後に最も高頻度に使われる前置詞を確認しなさい。

Q5. Advancedにて "research" と入力しPart of speechをverb、Text typesをSpokenと指定し、researchに続く最も高頻度に使われる前置詞を確認しなさい。

WORD SKETCH British National Corpus (BNC)

BASIC **ADVANCED** AS A LIST ABOUT

Search ? Subcorpus ?
research none (the whole corp...)

Part of speech ?

- auto
- noun
- verb**
- adjective
- adverb
- pronoun

Text types (1) ? ^

Text type ^

- Spoken context-go... X
- Spoken demographi... X**

Written books and periodicals

Written miscellaneous

Written-to-be-spoken



GO

prepositional phrases

"research" into ...	2.7%	...
"research" throughout ...	2.7%	...

Q6. Word Sketch: BNC利用

Basicにて "sign" を検索後、Part of speech をverbに変更し、sign の目的語として最もよく共起して使われる名詞を確認しなさい。

Q6. Basicにて "sign" を検索後、Part of speechをverbに変更し、"sign"の目的語として最もよく共起して使われる名詞を確認しなさい。

WORD SKETCH British National Corpus (BNC)

BASIC ADVANCED AS A LIST ABOUT

Search ?

sign

GO

WORD SKETCH British National Corpus (BNC)

sign as noun 10,528x verb 9,068x

Sorted by frequency

Part of speech	Count	Frequency
adjective 4x	272	6.9
adverb 1x	114	6.5

prepositional phrases

- "sign" of ...
- ... of "sign"
- ... for "sign"

WORD SKETCH British National Corpus (BNC)

sign as verb 9,068x

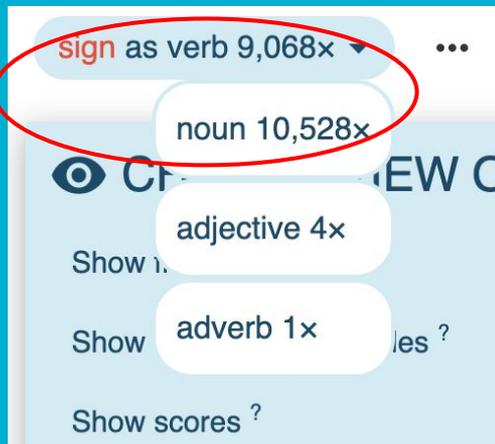
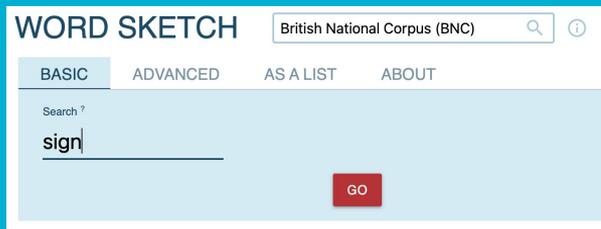
Sorted by frequency

Category	Item	Count	Frequency	
modifiers of "sign"	not	165	2.7	
	just	80	4.8	
	also	72	4.3	
	already	59	5.4	
	recently	37	6.6	
	then	32	4.1	
	now	27	3.5	
	yet	24	5.7	
	formally	23	7.7	
	actually	23	4.9	
	only	21	3.1	
	objects of "sign"	agreement	803	11.3
		contract	378	10.4
		treaty	308	10.7
document		164	9.4	
deal		151	8.8	
letter		146	8.6	
form		133	8.2	
petition		110	9.3	
declaration		72	8.7	
protocol		67	8.7	
accord		63	8.6	
name		58	6.8	
subjects of "sign"		people	52	4.9
		government	49	6.0
	country	41	7.0	
	party	34	6.3	
	person	32	6.5	
	corp	31	7.8	
	company	28	5.5	
	leader	26	7.2	
	member	26	6.4	
	inc	23	7.2	

Q7. Word Sketch: BNC利用

Basicにて "sign" を検索後、Part of speech をnounに変更し、"sign" を最も強い共起関係を持って目的語として取る動詞を確認しなさい。

Q7. Basicにて "sign" を検索後、Part of speechをnounに変更し、"sign" を最も強い共起関係を持って目的語として取る動詞を確認しなさい。



verbs with "sign" as object		
show	917	10.0
display	35	7.2
erect	20	6.8
warn	23	6.8
the warning signs		
detect	19	6.5
notice	19	6.4
read	34	6.4
read the signs		
be	2,172	6.4
see	216	6.4
spot	14	6.2
recognise	21	6.2
recognised the signs		
illuminate	12	6.1

Q8. Word Sketch Differenceの機能の概要を確認しなさい。

Get more space 



Q8. Word Sketch Differenceの機能の概要を確認しなさい。



 **Quick start guide**
Basics in a few minutes

 **User guide**
Everything about Sketch Engine

 **Ask a question**
Let our support team help you

 **Interface introduction**
Show 1-minute introduction to the interface

 **Get notified**
About new corpora and tools

-  [How to use the word sketch difference](#)
-  [What makes the word sketches difference unique?](#)
-  [How does it work?](#)
-  [Requirements for the word sketch difference to work well](#)
-  [Referencing word sketches, bibliography](#)

corpus tools

[word sketch](#) [sketch difference](#) [thesaurus](#) [concordance](#)
[parallel concordance](#) [wordlist](#) [n-grams](#) [keywords & terms](#)
[bilingual term extraction](#) [trends](#) [OneClick Dictionary](#)
[text type analysis](#)

Q9. Word Sketch Difference: BNC利用

First lemmaに"fat"、Second lemmaに"fatty"を入力した後、「nouns and verbs modified by "fat/fatty"」の検索結果を確認しなさい。

Show visualisation をクリックして「fat」と「fatty」とそれぞれ共起する語の特徴を見出しなさい。



BASIC ADVANCED ABOUT

First lemma ? Second lemma ?

fat fatty

GO

Q9. Word Sketch Difference: BNC利用



CHANGE VIEW OPTIONS

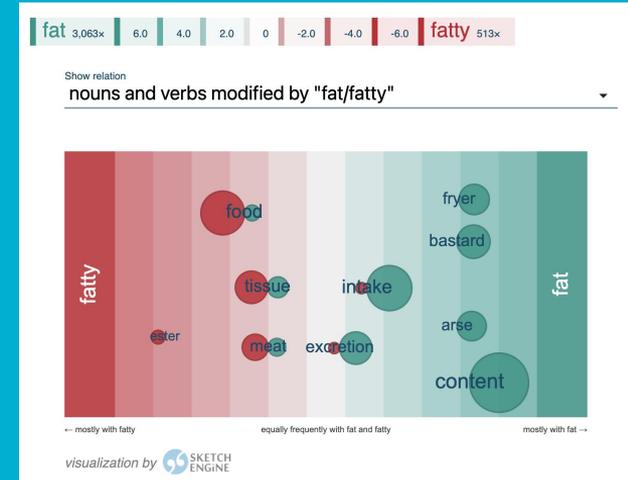
Show counts ?

Show scores ?

The typicality score ([LogDice](#)) of the second lemma or word form or in the second subcorpus. It indicates how strong the collocation is. The higher the score, the stronger the collocation. A low score means that the words in the collocation also frequently combine with many other words.

nouns and verbs modified by "fat/fatty"

content	8.5	—	...
fryer	7.9	—	...
arse	7.8	—	...
bastard	7.8	—	...
intake	8.5	5.4	...
excretion	7.9	6.2	...
meat	5.7	7.8	...
tissue	6.1	8.1	...
food	4.1	7.4	...
ester	—	7.4	...
biosynthesis	—	7.6	...
acid	—	10.9	...



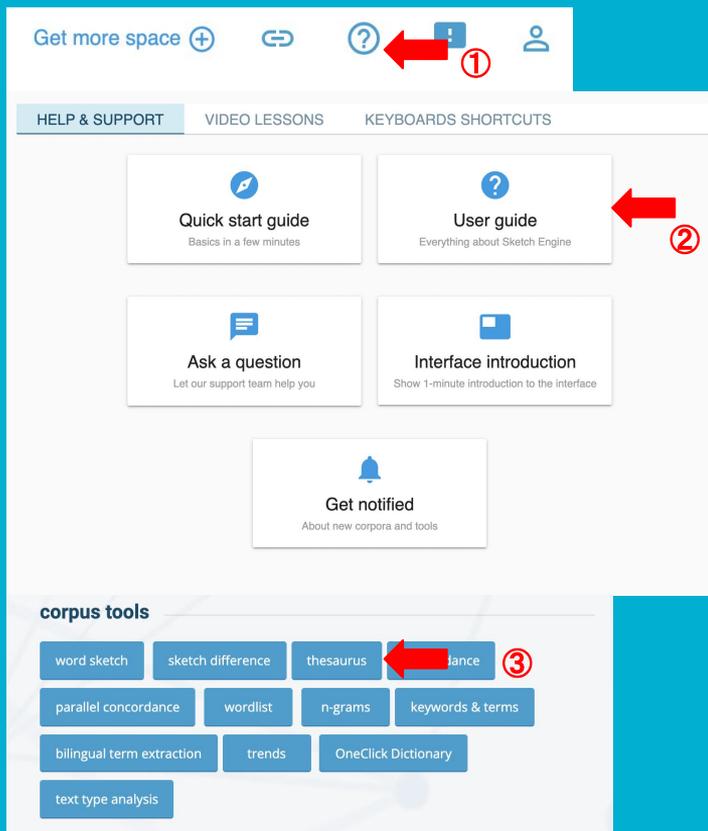
**Q10. Word Sketch Difference: BNC利用
Sketch Difference機能を使って"interesting" と
"interested"の使用傾向の違いを確認しなさい。**

Q11. Word Sketch Difference: BNC利用
Sketch Difference機能を使って "random" と "arbitrary"
の使用傾向の違いを確認しなさい。

Q12. Word Sketch Difference: BNC利用
Sketch Difference機能を使って"convincing"と
"persuasive"の類似点と相違点を見出さない。

**Q.13 ThesaurusのUser Guide
を概観して概要を把握しなさい。**

Q13. ThesaurusのUser Guideを概観して概要を把握しなさい。



Get more space +

HELP & SUPPORT VIDEO LESSONS KEYBOARDS SHORTCUTS

Quick start guide
Basics in a few minutes

User guide
Everything about Sketch Engine

Ask a question
Let our support team help you

Interface introduction
Show 1-minute introduction to the interface

Get notified
About new corpora and tools

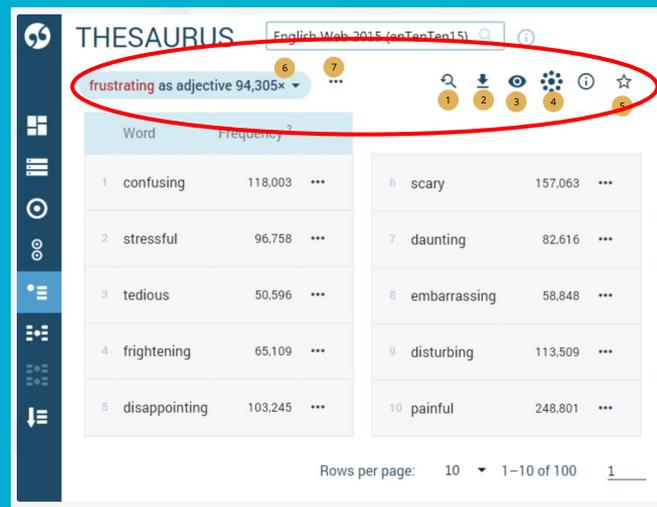
corpus tools

word sketch sketch difference thesaurus

parallel concordance wordlist n-grams keywords & terms

bilingual term extraction trends OneClick Dictionary

text type analysis

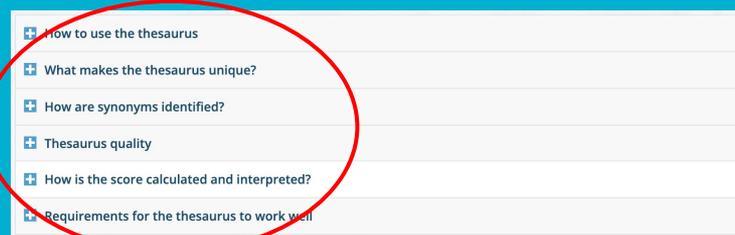


THESAURUS English Web 2015 (enTenTen15)

frustrating as adjective 94,305x

Word	Frequency
1 confusing	118,003
2 stressful	96,758
3 tedious	50,596
4 frightening	65,109
5 disappointing	103,245
6 scary	157,063
7 daunting	82,616
8 embarrassing	58,848
9 disturbing	113,509
10 painful	248,801

Rows per page: 10 1-10 of 100



- How to use the thesaurus
- What makes the thesaurus unique?
- How are synonyms identified?
- Thesaurus quality
- How is the score calculated and interpreted?
- Requirements for the thesaurus to work well

Q14. Thesaurusでbigを検索し、
どのような単語があらわれるか確
認しなさい。

Q14. Thesaurusでbigを検索し、どのような単語があらわれるか確認しなさい。

Sketch EngineのThesaurus機能では、Similarityスコアに基づいて語を抽出しています。

これは共起する語が検索語と似ている語を評価します。つまり検索語と似たような環境で使われる語を抽出することになります。語の意味に基づいて語を抽出しているわけではないので、意味を考えると類語とは言えないものも抽出されます。bigを検索すると2位にsmallがくるところからわかるように、対義語が抽出されることも多くあります。

したがって、純粋にThesaurusというよりかは、似たような環境で使われる語(共起語)が似ている語を抽出する機能であると考えた方が实际的です。

	Word	Frequency [?]	Similarity [?] ↓
1	large	47,314	0.524 ...
2	small	50,758	0.492 ...
3	huge	7,631	0.442 ...
4	old	57,937	0.431 ...
5	great	59,408	0.427 ...
6	new	105,645	0.420 ...
7	little	48,172	0.404 ...
8	few	46,390	0.396 ...
9	own	67,022	0.392 ...
10	second	31,784	0.379 ...

10 ▼ 1-10 of 1,000 ⏪ < 1

Q15. Concordanceで”provide”を、活用形も含めて抽出し
KWIC表示しなさい。

Q15. Concordanceで”provide”を、活用形も含めて抽出しKWIC表示しなさい。

ConcordanceのBASICタブで単語を打ち込めば、その単語をレマ検索でき、活用形も含めたものを抽出することができます。

provideと打ち込めば、provideの活用形も含めて検索し、コンコーダンス上でKWIC表示することができます。KWIC表示になっていない場合は、**右上の青いタブでSentence表示とKWIC表示を切り替えることができます。**

ADVANCEDタブのQuery type: simpleでも同じように検索ができます。

The screenshot shows the British National Corpus (BNC) Concordance tool. The search term is 'provide', with 54,620 results. The interface is in the 'BASIC' tab. The search results are displayed in a table with columns for 'Details', 'Left context', 'KWIC', and 'Right context'. A red arrow points to the 'KWIC' tab in the top right corner of the interface.

Details	Left context	KWIC	Right context
1	Written books a... little assistance . . .	provide	a richer source of information . . .
2	Written books a... rops or craft areas . . .	provide	evidence for only a small number of mosaics are included . . .
3	Written books a... certain aspects of it which are noteworthy throughout . . .	provide	a useful introduction to the elements of stylistic comparison : (i) The ur . . .
4	Written books a... e distances , or was altered by heat , but most major sources were local and . . .	provided	stone which was used in its natural state . . .
5	Written books a... rth Leigh , the quarries of Stonesfield are only one mile distant and will have . . .	provided	the bulk of the friable stone , while the quarries of oolite at Stuart wood (les . . .
6	Written books a... e , while the quarries of oolite at Stuart wood (less than 200 metres away) . . .	provided	that of higher quality . . .
7	Written books a... arent width of black and white lines had to be carefully adjusted in order to . . .	provide	a convincing effect . . .
8	Written books a... s><s> 71) , was almost certainly planned as a whole but , nevertheless , . . .	provides	a useful illustration of these considerations . . .
9	Written books a... Som , (Arch . . .	provides	the best examples . . .
10	Written books a... ars further corroboration : certain sizes of tesserae and motifs are shown to . . .	provide	maximum adaptability of design , because they are proportional to the abstr . . .
11	Written books a... see Appendix E , where a list of the pavements under discussion is also . . .	provided	. . .
12	Written books a... y of its mosaics are very early , and so outside the scope of this paper , it . . .	provides	invaluable insights into the origins and development of British mosaics at a . . .
13	Written books a... perhaps of long-term developments of culture and of changes in taste , is . . .	provided	by a consideration of the use of structure and of some motifs in the compo . . .
14	Written books a... employed in a similar " framing " role . . .	provides	an outer border for the Winchester mosaic . . .
15	Written books a... other than insets which would be subsumed by the whole design rather than . . .	provide	an independent pictorial element . . .
16	Written books a... c link between mosaic A and the Cupids mosaic . . .	provides	a very interesting suggestion as to the direction and strength of co-operati . . .
17	Written books a... s the quality and effect of the design in the central panel , however , which . . .	provides	the closest parallel with the Middleborough mosaic . . .
18	Written books a... om room 30 , North Leigh . . .	provide	almost identical versions of the same figural scene (pls. 13 and 14) . . .

検索結果

Q16. Concordanceで名詞の”play”を、複数形も含めて検索
しなさい。

Q16. Concordanceで名詞の”play”を、複数形も含めて検索しなさい。

BASICタブでは品詞を指定できないため、”play”を検索すると、動詞のplayも名詞もplayも混ざって抽出されてしまいます。

品詞を指定した検索は、ADVANCEDタブのQuery type: lemmaで行うことができます。Query typeのlemmaを選択すると、品詞が選択できるので、nounを選択し、”play”と打ち込んで検索します。

The screenshot shows the 'CONCORDANCE' search interface for the British National Corpus (BNC). The 'ADVANCED' tab is active. The 'Query type' dropdown is set to 'lemma', and the 'Part of speech' dropdown is set to 'noun'. The search term 'play' is entered in the search box. The 'Subcorpus' is set to 'none (the whole corpus)' and the 'Macro' is set to 'none'. A red 'GO' button is visible at the bottom right.

Q17. Concordanceで“dis-”から始まる単語を検索しなさい。

Q17. Concordanceで“dis-”から始まる単語を検索しなさい。

ADVANCEDタブのsimple検索で、dis* と打ち込むとdisから始まる単語を抽出することができます。

simple検索では*(アスタリスク)が使用でき、dis*と検索すればdisから始まる全ての単語(活用形も含む)が検索でき、*ingと検索すれば、ingで終わる全ての単語を検索することができます。

CONCORDANCE British National Corpus (BNC)

simple dis* • 267,272
2,379.01 per million tokens • 0.24%

Details KWIC Right context

Line	Left context	Word	Right context
1	Written books a... considering Style in Roman Society </s><s> It is notoriously difficult to	discuss	the role of aesthetics in Rome and to ascribe definite aesthetic interests to
2	Written books a... ist . </s><s> Nat . 30.60.64) , meanwhile , echoes much of that	discussed	by Vitruvius - with only slight differences , e.g. the types of pavements to
3	Written books a... d under " form " , in section 2.3 (all section numbers refer to the under	discussion	, unless otherwise indicated) . </s><s> Between mosaics only one oi
4	Written books a... s is relatable only by comparison with that of a third pavement (a "	distributed	relationship) . </s><s> (iv) Unitary . </s><s> At this level of .
5	Written books a... ique to the style are conspicuous . </s><s> It is easy to blur such	distinctions	. </s><s> It is especially easy to confuse unitary with integral affi
6	Written books a... nited in number . </s><s> Some stone was transported considerable	distances	, or was altered by heat , but most major sources were local and provided
7	Written books a... </s><s> At North Leigh , the quarries of Stonesfield are only one mile	distant	and will have provided the bulk of the friable stone , while the quarries of c
8	Written books a... /s><s> However , sites were not usually so well located , and even when	distant	from sources of suitable stone , there can have been little special movme
9	Written books a... l in the pavement of room 2 . </s><s> The few instances of " long	distance	transport tend to be early : at Fishbourne , or Angmering , where siltstor
10	Written books a... nces in the Charioteer mosaic , from Rudston , E . Yorks . , (now	displayed	in the Museum of Transport and Archaeology , Hull) . </s><s> Of thi
11	Written books a... Westland , Som. , for example , tesserae are said to have fallen into three	distinctive	sizes : 1 cm . sq . ; 2.5 cm . sq . ; and 3.5 cm . sq .
12	Written books a... /s><s> As lighter shades of tesserae appear nearer when viewed from a	distance	, so do larger tesserae ; conversely , small tesserae appear more distant
13	Written books a... stance , so do larger tesserae ; conversely , small tesserae appear more	distant	, as do darker shades . </s><s> Such considerations are , obviously
14	Written books a... much so that at Woodchester Orpheus may , for this reason , have been	displaced	from his usual , central position (pi . </s><s> 43 ; although see pi
15	Written books a... section . </s><s> 3.6 Aids to Design </s><s> Guidelines have been	discovered	beneath British mosaics on only a few occasions : the author knows of or
16	Written books a...) - although , in wall painting such practices were also known (plaster on	display	, Jewry Wall Museum) . </s><s> The complexity of many designs , f
17	Written books a... p is implied . </s><s> When speaking of laying tesserae , the main	distinction	to be made is between direct and indirect methods of application : the c

検索結果

CONCORDANCE British National Corpus (BNC)

BASIC ADVANCED ABOUT

Query type

Simple
dis*

lemma
phrase
word
character
CQL

Q18. BNCで、ofとimportanceの間にくる単語のうち、レマ化した頻度で第5位を調べなさい。

Q18. BNCで、ofとimportanceの間にくる単語のうち、レマ化した頻度で第5位を調べなさい

paramountが正解です。

ADVANCEDタブのsimple検索で下記検索画面のように検索します。simple検索において、*(アスタリスク)は前後にスペースを入れて打ち込むと、任意の1語を意味します。

検索結果を頻度順リストにして表示するには、右上のFREQUENCYを選びます。BASICタブでKWICのLEMMAを選びます。WORD FORMSだと表記形(活用形や大文字・小文字が別々にカウントされる)、LEMMAだとレマ化して抽出されるので多少結果が異なります。

CONCORDANCE British National Corpus (BNC)

simple of * importance 2,512
22.36 per million tokens • 0.0022%

Details Left context KWIC Right context

1	Written books a...	3.2 The colour of Tesseræe	</s><s> Sensitive use of shading is	of obvious importance	- to appreciate this one could little better than study the strik
2	Written books a...	irker shades	</s><s> Such considerations are, obviously,	of vital importance	in figural mosaics which hope to maintain a degree of naturalisr
3	Written books a...	</s><s> It is	the design of room II, however, which is	of special importance	</s><s> The emphasis of such arrangements is the most

FREQUENCY

BASIC ADVANCED ABOUT

First word to the left KWIC First word to the right

WORD FORMS WORD FORMS WORD FORMS

PART OF SPEECH PART OF SPEECH PART OF SPEECH

TAGS TAGS TAGS

LEMMA LEMMA LEMMA

More presets

TEXT TYPES

LINE DETAILS

BASIC ADVANCED ABOUT

Query type

simple

lemma

phrase

word

(218 items, 2,512 total frequency)

	Lemma	Frequency ↓	Relative ?	
1	of the importance	392	3.49	...
2	of great importance	249	2.22	...
3	of particular importance	131	1.17	...
4	of crucial importance	112	1.00	...
5	of paramount importance	109	0.97	...
6	of considerable importance	87	0.77	...
7	of vital importance	79	0.70	...
8	of some importance	72	0.64	...
9	of prime importance	70	0.62	...
10	of central importance	59	0.53	...

Q19 名詞のincreaseの1つ後ろにくる語と3つ後ろの来る語の頻度を調べ、レマ単位でそれぞれ第1位は何か確かめなさい。

Q20. account forというフレーズを使った、学習者に提示すべき例文をコンコードダンスを参考にして考えなさい。

Q20. account forというフレーズを使った、学習者に提示すべき例文をコンコーダンスを参考にして考えなさい。

コンコーダンスの機能の一つに、Good dictionary expmplesというものがあります。

この機能を使うと、検索した表現を含む文を、一定の基準に従って、例文として分かりやすいと思われる順に並べてくれます。

この中から例文にふさわしいと思うものを選べば、効率よく例文を作ることができます。

The screenshot shows the CONCORDANCE interface for the British National Corpus (BNC). The search term is 'simple account for', with 4,221 results and 0.0038% of the corpus. The interface includes a search bar, a toolbar with various icons, and a list of search results. A red arrow points to the 'Good dictionary expmples' icon in the toolbar. The search results are sorted by GDEX and show the following examples:

Result ID	Text
1	<S> No share premium has been accounted for on this issue . </S>
2	<S> It would account for the strong resemblance . </S>
3	<S> They now account for forty per cent of sales . </S>
4	<S> He wants every single penny accounted for before he starts doing anything . </S>
5	<S> At least two reasons account for this, besides general confidence . </S>
6	<S> Money was being spent, but money could be accounted for . </S>
7	<S> Their numbers are known and supposedly accounted for by the vergers . </S>
8	<S> Between them, these bacteria account for many losses suffered by fishkeepers . </S>
9	<S> He did, which accounts for the show's huge success . </S>
10	<S> Open a bank account for yourself . </S>
11	<S> What accounts for the fatal flaws ? </S>

Q21. 動詞の原形のprocessをコンコーダンスに表示しなさい。

Q21. 動詞の原形のprocessをコンコーダンスに表示しない。

Query type: wordを選ぶと、打ち込んだ形のもののみが抽出されます。processと検索すれば、活用形は含まず、playのみが抽出されますし、processesと検索すれば、processesのみが抽出されます。

The screenshot shows the CONCORDANCE search interface for the British National Corpus (BNC). The interface is divided into three tabs: BASIC, ADVANCED, and ABOUT. The ADVANCED tab is selected. The 'Query type' dropdown is set to 'word', and the 'Part of speech' dropdown is set to 'verb'. The search term 'process' is entered in the search field. The 'GO' button is visible at the bottom right.

Q22. Concordanceで“spect”を含む語を検索しなさい。

Q22. Concordanceで“spect”を含む語を検索しなさい。

Query typeのcharacter検索では、入力した文字列を含む語を検索できます。

spectと入力すれば、aspect, respect, prospect, spectatorなどが抽出されます。

The screenshot shows the 'ADVANCED' tab of a concordance search interface. The 'Query type' dropdown menu is open, with 'character' selected and circled in red. The search term 'spect' is entered in the 'Character' field. Other options include 'simple', 'lemma', 'phrase', 'word', and 'SQL'. The 'Subcorpus' is set to 'none (the whole corpus)' and the 'Macro' is set to 'none'. There are also 'Filter context' and 'Text types' dropdowns, and a red 'GO' button.

Q23. Parallel Concordance:

パラレルコンコーダンスで、JEFLIを用い、makeを検索しなさい

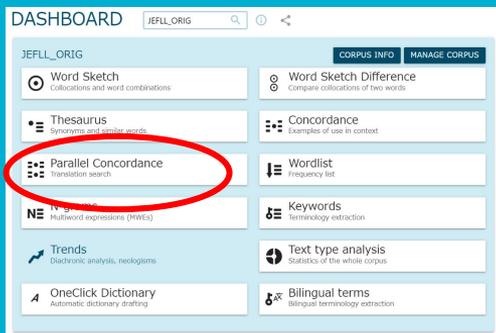
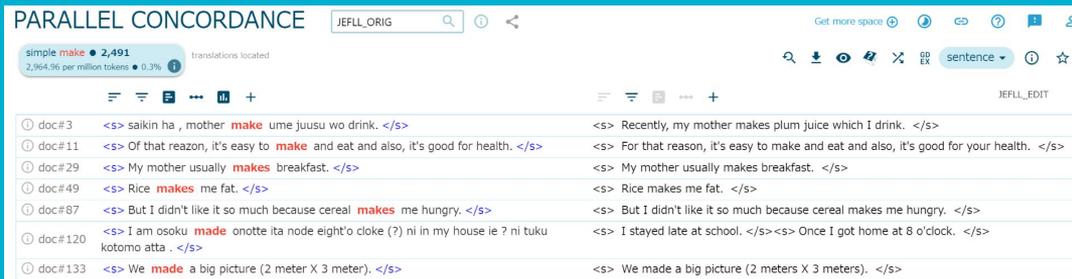
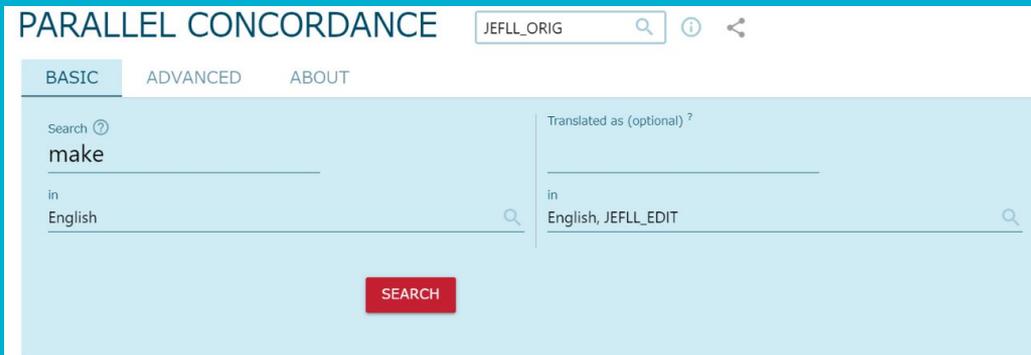
Q23. パラレルコンコーダンスで、JEFLをうい、makeを検索しなさい

東京外国語大学内でのアカウントあれば JEFL、NICER、ICNALE Edited Essaysの3種の学習者コーパスにおいて、パラレルコンコーダンスを利用できます。

学習者コーパスのパラレルコンコーダンスでは、学習者が産出した英文とそれに対応する母語話者による添削文を同時に表示することができます。

BASICタブで、左側に検索したい語を入力します。右側には何も入力する必要はありません。

コンコーダンスでは、左側に学習者の文、右側にそれに対応する添削文が表示されます。



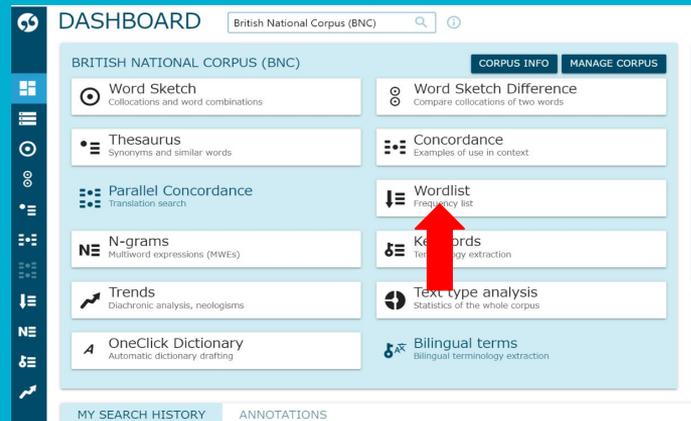
Q24. Wordlist:
BNCで最もよく使われる単語を
確認しなさい。

Q24. Wordlist: BNCで最もよく使われる単語を確認しなさい。

Wordlistの機能を使うと、簡単に語の頻度リストを出すことができます。

BASICタブで、何もせずそのままGOをクリックすると、コーパス全体での語彙頻度リストを出すことができます。

最もよく使われる単語はtheであることが簡単に分かります。



word (165,953 items | 111,680,004 total frequency)

Word	Frequency [?] ↓	Frequency Per Million [?] ↓	DOCF [?]
1 the	6,054,939	53,895.59	4,050 ...
2 ,	5,063,028	45,066.50	4,045 ...
3 .	4,818,451	42,889.49	4,054 ...
4 of	3,049,448	27,143.43	4,040 ...
5 and	2,624,147	23,357.78	4,050 ...
6 to	2,599,451	23,137.96	4,049 ...
7 a	2,175,967	19,368.49	4,045 ...
8 in	1,945,533	17,317.38	4,047 ...
9 '	1,562,877	13,911.32	3,628 ...

Q25. Wordlist:
BNCで副詞の頻度リストを出しな
さい。

Q25. BNC全体で、副詞の頻度リストを出しなさい

Wordlistの機能では、品詞を指定せず表記形(words)で抽出したり、レマ単位(lemmas)で抽出したり、品詞を指定して抽出したりすることができます。

副詞を指定するには、adverbを選択してリストを出します。

WORDLIST British National Corpus (BNC)

BASIC ADVANCED ABOUT

find ? words lemmas adjective **adverb** conjunction noun preposition pronoun

all starting with ending with containing

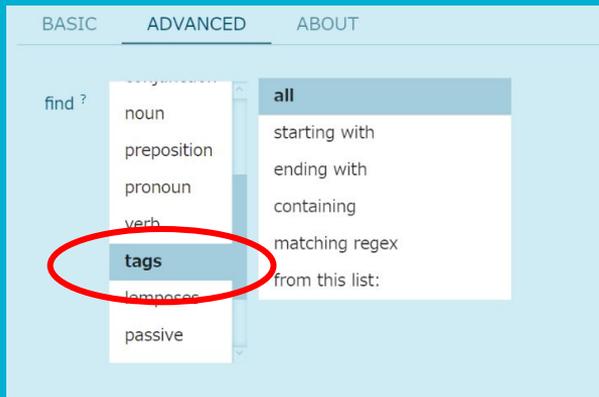
GO

Q26. Wordlist:
BNCで最もよく使われる品詞を
確認しなさい。

Q26. BNCで最もよく使われる品詞を確認しなさい。

単語ではなく品詞単位で頻度リストを出す場合は、tagsを選びます。

すると右のようにタグの頻度表を出すことができます。NNは単数形の名詞を表すタグであり、名詞が最も使われる品詞であることが分かります。



tag (56 items 103,764,845 total frequency)			
Tag	Frequency [?] ↓	Frequency Per Million [?] ↓	DOCF [?]
1 NN	14,752,903	131,317.00	4,054 ...
2 IN	12,512,020	111,370.68	4,054 ...
3 DT	10,227,865	91,039.20	4,054 ...
4 JJ	7,009,866	62,395.49	4,052 ...
5 NP	6,564,177	58,428.37	4,053 ...
6 NNS	5,347,278	47,596.63	4,049 ...
7 SENT	5,345,332	47,579.31	4,054 ...
8 RB	5,294,068	47,123.00	4,053 ...
9 PP	5,101,166	45,405.97	4,051 ...
10 CC	3,498,280	31,138.52	4,050 ...
11 VV	2,884,680	25,676.81	4,053 ...
12 VVN	2,662,832	23,702.12	4,050 ...
13 VVD	2,240,920	19,946.64	4,022 ...
14 CD	1,977,975	17,606.14	4,035 ...
15 VVG	1,725,050	15,354.83	4,046 ...
16 TO	1,597,628	14,220.64	4,045 ...
17 MD	1,411,644	12,565.18	4,047 ...
18 PPZ	1,409,171	12,543.17	4,039 ...
19 VBZ	1,305,154	11,617.30	4,045 ...
20 VBD	1,196,648	10,651.48	3,958 ...

Q27. N-gram: BNC利用

Basic検索で「4」を指定して、4語の羅列で最も使われる表現が "I don't know"であることを確認しなさい。

Q27. N-gram: BNC利用

Basic検索で4を指定して、4語の羅列で最も使われる表現が "I don't know"であることを確認しなさい。

N-GRAMS

BASIC ADVANCED ABOUT

N-gram length ? 2 3 4 5 6

GO



	Word	Frequency ?
1	i do n't know	11,921 ...
2	the end of the	10,591 ...
3	at the end of	9,375 ...
4	i do n't think	6,996 ...
5	at the same time	6,924 ...
6	for the first time	5,438 ...
7	on the other hand	5,311 ...
8	as a result of	5,156 ...
9	the rest of the	5,155 ...
10	in the case of	4,723 ...

Q28. N-gram: BNC利用

Advanced 検索で「4」を指定し、Attributeを lemma、Additional criteria を all、A = aをチェックした後、Subcorpus を Witten Academic に変更し検索しなさい。

Q28. N-gram: BNC利用

Advanced 検索で「4」を指定し、Attributeをlemma、Additional criteria をall、A = aをチェックした後、Subcorpus を「Written Academic」に変更し検索しなさい

CHANGE CRITERIA

BASIC **ADVANCED** ABOUT

N-gram length ? 2 3 **4** 6

Attribute ?
lemma

Frequency min ? 5 Frequency max ? 0

Subcorpus ?
Written Academic

Nest n-grams ? Include nonwords ?

A = a ? Exclude these words: ?

Key n-grams ?

Additional criteria ?

- all**
- starting with letters
- ending with letters
- containing letters
- starting with word
- containing word
- ending with word
- matching regular expression

Subcorpus and Text types cannot be combined

	Word	Frequency ?
1	in the case of	2,084 ...
2	on the other hand	1,784 ...
3	the end of the	1,592 ...
4	at the same time	1,481 ...
5	as a result of	1,440 ...
6	the way in which	1,286 ...
7	the court of appeal	1,256 ...
8	on the basis of	1,163 ...
9	at the end of	1,149 ...
10	per cent of the	1,047 ...

Q29. N-gram: BNC利用

Advanced 検索

length「3」

Attribute「word」

Additional criteria「starting with word」

‘take’と入力

A = a をチェックして検索後、
Attributeを「lemma」に変更すると結果が変化することを確認しなさい。

Q29. N-gram: BNC利用

CHANGE CRITERIA

BASIC ADVANCED ABOUT

N-gram length? 2 3 4 5 6

Attribute? word

Frequency min? 5 Frequency max? 0

Subcorpus? none (the whole corp...)

Additional criteria

- all
- starting with letters
- ending with letters
- containing letters
- starting with word**
- containing word
- ending with word
- match a regular expression

Text types?

GO

Word	Frequency?
take advantage of	1,074 ...
2 take part in	1,060 ...
3 taken into account	1,040 ...
4 take account of	847 ...
5 take into account	841 ...
6 take place in	631 ...
7 taken by the	597 ...
8 taken from the	541 ...
9 take care of	539 ...
10 take on the	464 ...

Word	Frequency?
take into account	2,409 ...
2 take place in	2,391 ...
3 take part in	2,378 ...
4 take advantage of	1,728 ...
5 take over the	1,264 ...
6 take account of	1,223 ...
7 take on the	1,019 ...
8 take up the	931 ...
9 take care of	914 ...
10 take to the	890 ...

word 検索

lemma 検索

Q30. N-gram: BNC利用

n-gramの機能を利用してもっともよく使われる3語の品詞連鎖を確認しなさい。

Q30. n-gramの機能を利用してもっともよく使われる3語の品詞連鎖を確認しなさい。

Attributeをtagにすることでn-gram検索では品詞タグの連鎖を出すことができます。

結果を見るとN DT NN (前置詞＋限定詞＋名詞単数の連鎖)が最も多いことがわかります。

BASIC **ADVANCED** ABOUT

N-gram length ? 2 3 4 5 6

Attribute ?
tag

Nest n-grams ? Include nonwords ?

A = a ? Exclude these words: ?

Frequency min ? Frequency max ? Subcorpus ?

5 0 none (the whole corpus) +

Key n-grams ?

tag (items: 37,120 , total frequency: 88,734,686)

	Word	Frequency ?	Word	Frequency ?	
1	IN DT NN	2,227,366 ...	14	DT NN NN	416,979 ...
2	DT NN IN	1,720,758 ...	15	NN IN NP	399,266 ...
3	DT JJ NN	1,538,602 ...	16	IN DT NNS	391,535 ...
4	NN IN DT	1,490,502 ...	17	IN JJ NN	362,025 ...
5	IN DT JJ	1,019,304 ...	18	NN IN JJ	360,951 ...
6	JJ NN IN	881,357 ...	19	IN NN IN	360,454 ...
7	NN IN NN	651,954 ...	20	JJ NNS IN	360,232 ...
8	DT NN SENT	501,365 ...	21	IN JJ NNS	354,611 ...
9	NNS IN DT	500,263 ...	22	SENT DT NN	329,248 ...
10	VVN IN DT	465,972 ...	23	TO VV DT	324,874 ...
11	IN DT NP	440,563 ...	24	IN PPZ NN	324,598 ...
12	JJ NN SENT	437,145 ...	25	VV DT NN	320,733 ...
13	NP NP NP	428,409 ...	26	NN SENT DT	313,312 ...

Q31. 特徴語抽出: JEFLL_ORIG
をBAWE(British Academic
Written English Corpus)と比較
して最も特徴的な語がJPであるこ
とを確認しなさい。

Q31. 特徴語抽出: JEFLL_ORIGをBAWE(British Academic Written English Corpus)と比較して最も特徴的な語がJPであることを確認しなさい。

DASHBOARD JEFLL_ORIG

JEFLL_ORIG

- Word Sketch
- Word Sketch Difference
- Thesaurus
- Concordance
- Parallel Concordance
- Wordlist
- N-grams
- Keywords
- Trends
- Text type analysis
- OneClick Dictionary
- Bilingual terms

BASIC **ADVANCED** ABOUT

Focus subcorpus ?
none (the whole corp...)

Reference corpus ?
British Academic Written English C...

Reference subcorpus ?
none (the whole corpus)

Focus on ?
rare 1 common

Minimum frequency ? 100
Maximum frequency ? 0

Maximum items ? 1000

A = a ?
 At least one alphanumeric ?
 Only alphanumeric ?
 Include nonwords ?
 Exclude these words : ?
 From list

Identify keywords
 Identify terms
 Identify n-grams

Keywords settings
Attribute ?
lemma

Matching regex ?
.*

Terms settings
Matching regex ?
.*

N-grams settings
Attribute ?
word

Matching regex ?
.*

N-gram length 2 3 4 5 6

Word		
1	jp	...
2	otoshidama	...
3	urashima	...
4	them	...
5	taro	...
6	him	...
7	me	...
8	breakfast	...
9	ll	...
10	and_so_on	...

左上のアイコンをクリックすると、対象コーパスと参照コーパスを入れ替えて検索することができる。

SINGLE-WORDS ✓

↔ reference corpus: British Academic Written English Corpus (BAWE) (items: 530)

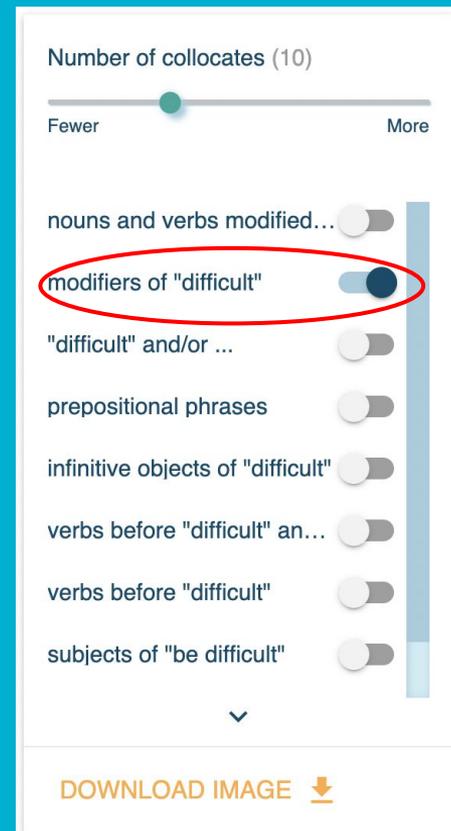
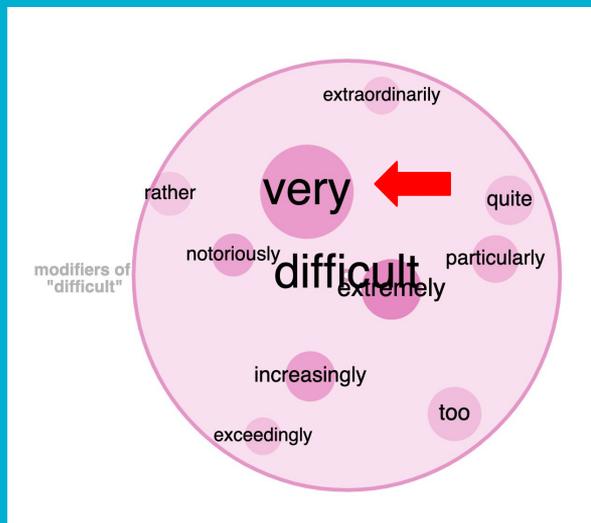
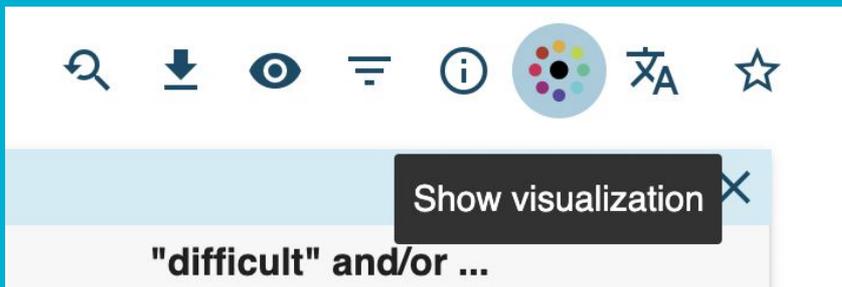
Word	Word	Word
1 jp ...	11 festival ...	21 baseball ...
2 otoshidama ...	12 for_example ...	22 bread ...
3 urashima ...	13 kame ...	23 お味噌汁 ...
4 them ...	14 miso ...	24 宝物 ...
5 taro ...	15 ryugujo ...	25 proper_noun ...
6 him ...	16 otohime ...	26 ryugujoyo ...
7 me ...	17 urashima-taro ...	27 soccer ...
8 breakfast ...	18 劇 ...	28 tamatebako ...
9 ll ...	19 otohime-sama ...	29 tenji ...
10 and_so_on ...	20 展示 ...	30 of_course ...

Rows per page: 50 1-50 of 506

中級編

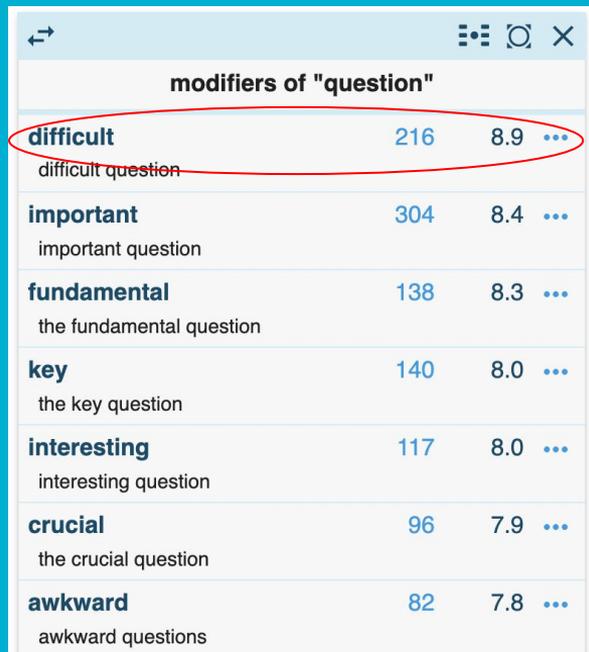
Q32. Word Sketch: BNC利用
Basicにて"difficult" を検索し、Show
visualization 機能を用いてdifficultを修飾す
る語 (modifier)として最も高頻度に使われる
語を確認しなさい。

Q32. Basicにて"difficult" を検索し、Show visualization 機能を用いてdifficultを修飾する語 (modifier)として最も高頻度に使われる語を確認しなさい。



Q33. Word Sketch: BNC利用
Basicにて"question"を検索し、"question"
と最も強く共起する修飾語が "difficult"であ
ることを確認しなさい。

Q33. Basicにて"question"を検索し、"question"と最も強く共起する修飾語が "difficult"であることを確認しなさい。



The screenshot shows a table titled "modifiers of 'question'". The table lists various adjectives used to modify the word "question", along with their frequency and a score. The "difficult" row is circled in red.

Modifier	Count	Score	Example
difficult	216	8.9	difficult question
important	304	8.4	important question
fundamental	138	8.3	the fundamental question
key	140	8.0	the key question
interesting	117	8.0	interesting question
crucial	96	7.9	the crucial question
awkward	82	7.8	awkward questions

Q34. Word Sketch: BNC利用
difficult + question をMultiword Sketch機能を使ってWord Sketch検索し、"difficult question"を最も強い共起関係で目的語として取る動詞は"pose"であることを確認しなさい。

Q34. difficult + question をMultiword Sketch機能を使ってWord Sketch検索し、"difficult question"を最も強い共起関係で目的語として取る動詞は"pose"であることを確認しなさい。

modifiers of "question"		
difficult	216	8.9
difficult question		
important	4	
important question		
fundamental	138	8.3
the fundamental question		
key	140	8.0
the key question		

Context menu for "important":

- question + difficult
- question + difficult** (selected)
- difficult

verbs with "difficult question" as object		
pose	7	6.9
tackle	2	5.0
raise	11	4.6
ask	12	4.5
answer	2	4.1
avoid	2	3.2
become	2	1.3
be	61	1.2

Q35. Word Sketch: BNC 利用
"seize control"は固定表現でしょう
か？ "control"以外何が入るか
確認しなさい。

Q35. Word Sketch: BNC 利用

"seize control"は固定表現でしょうか？ "control"以外何が入るか確認しなさい。

CHANGE CRITERIA

BASIC ADVANCED AS A LIST ABOUT

Search ? seize

Subcorpus ? none (the whole corp...) 🔒 +

Part of speech ?

- auto
- noun
- verb**
- adjective
- adverb
- pronoun

Minimum frequency ? auto

Minimum score ? 0

Translate ?

Text types ? ▾

GO

objects of "seize"

opportunity	166	9.1	...
seized the opportunity			
initiative	50	8.9	...
to seize the initiative			
power	112	8.1	...
seized power in			
throne	15	8.0	...
wrist	15	7.9	...
control	55	7.8	...
seized control of the			
chance	58	7.6	...
and	33	7.5	...
asset	19	7.5	...
weapon	15	7.1	...
cocaine	7	7.0	...
pen	9	6.9	...

Q36. Word Sketch: BNC 利用
"put down"の間に入る名詞を確認
しなさい。"put down"の後に続
く名詞も確認しなさい。

Q36. Word Sketch: BNC 利用

"put down"の間に入る名詞を確認しなさい。"put down"の後に続く名詞も確認しなさい。

BASIC ADVANCED AS A LIST

Search ?

put

Part of speech ?

- auto
- noun
- verb**
- adjective
- adverb
- pronoun

particles after "put" with object

down put down	put + down
on put it on	Multiword Sketch put + down
in put it in	down
up put up	1,667 9.6 ...
away put away	272 9.5 ...
out put out	1,292 9.5 ...
off	555 9.3 ...
aside put aside	117 9.0 ...

objects of "put ... down"

foot	23	7.2 ...
glass	11	6.7 ...
head	23	5.7 ...
name	12	5.1 ...
hand	9	4.3 ...

objects of "put ... down down"

phone	69	11.4 ...
cup	37	10.6 ...
receiver	36	10.5 ...
root	30	10.3 ...
glass	27	10.1 ...
knife	22	9.8 ...
telephone	16	9.4 ...
book	15	9.3 ...
pen	15	9.3 ...
bag	13	9.1 ...
paper	9	8.6 ...
deposit	9	8.6 ...

Q37. Word Sketch: BNC利用

"naked"に続く名詞は全て身体の一部を表す語でしょうか？身体を表す単語以外何が使われるか確認しなさい。

Q37. "naked"に続く名詞は全て身体の一部を表す語でしょうか？身体を表す単語以外何が使われるか確認しなさい。

WORD SKETCH British National Corpus (BNC)

BASIC ADVANCED AS A LIST ABOUT

Search ?

naked

nouns and verbs modified by "naked"

breast	24	9.0	...
her naked breasts			
eye	143	8.7	...
with the naked eye			
chest	15	8.2	...
naked chest			
bulb	12	7.9	...
naked bulb			
shoulder	14	7.9	...
naked shoulders			
corpse	7	7.8	...
body	72	7.6	...
naked body			
aggression	8	7.6	...
ape	6	7.4	...
lunch	12	7.4	...
the naked lunch			
torso	5	7.3	...
flame	7	7.3	...

Q38. Word Sketch:

ドナルド・トランプと共起する形容詞

・名詞にはどんな物があるでしょうか？

English Web 2020 (enTenTen20)

コーパスを使って確認しなさい。

Q38. ドナルド・トランプと共起する形容詞・名詞にはどんな物があるでしょうか？

WORD SKETCH English Web 2020 (enTenTen20) 🔍 ⓘ

BASIC ADVANCED AS A LIST ABOUT

Search ?

Donald Trump

GO

↔ ⋮ 🔍 ×

adjective predicates of "Donald Trump"

unfit	7.9 ...
unqualified	6.9 ...
unpopular	6.6 ...
racist	6.2 ...
lookalike	6.0 ...
apocalyptic	5.7 ...
fit	5.7 ...
divisive	5.6 ...
presidential	5.6 ...
infamous	5.5 ...
reckless	4.8 ...
recent	4.8 ...

↔ ⋮ 🔍 ×

"Donald Trump" is a ...

nominee	8.5 ...
racist	8.5 ...
president	8.4 ...
president-elect	8.4 ...
apprentice	7.6 ...
President	7.5 ...
businessman	7.3 ...
supremacist	7.1 ...
Hitler	7.1 ...
candidate	7.0 ...
liar	6.7 ...
bigot	6.7 ...

Q39. Word Sketch Difference: BNC利用
"co-operate"と "collaborate"を検索し前置詞 (in, on, with, for, to, etc.)に続く語の変化を確認しなさい。

Q39.

"co-operate/collaborate" in ...					"co-operate/collaborate" to ...				
investigation	6	0	9.0	— ...	extent	3	0	5.4	— ...
venture	3	0	8.7	— ...					
coup	2	0	8.3	— ...					
making	2	0	7.5	— ...					
treatment	2	0	7.0	— ...					
effort	4	2	7.2	6.3 ...					
project	2	2	6.2	6.3 ...					
research	2	3	6.2	6.8 ...					
design	0	2	—	6.2 ...					
preparation	0	2	—	6.6 ...					
pair	0	2	—	7.4 ...					
delivery	0	2	—	8.6 ...					

"co-operate/collaborate" on ...					"co-operate/collaborate" for ...					"co-operate/collaborate" with ...				
technology	3	0	8.6	— ...	year	0	3	—	3.0 ...	investigation	9	0	9.5	— ...
basis	2	0	3.5	— ...						police	11	0	9.0	— ...
development	4	4	7.5	7.4 ...						committee	4	0	8.2	— ...
project	4	9	7.4	8.5 ...						officer	5	0	8.2	— ...
issue	0	2	—	4.8 ...						government	12	2	8.4	5.9 ...
work	0	2	—	6.1 ...						authority	15	3	8.8	6.5 ...
research	0	2	—	7.2 ...						teacher	3	6	7.2	8.4 ...
range	0	3	—	7.5 ...						stb	0	2	—	8.0 ...
book	0	3	—	7.5 ...						artist	0	3	—	8.2 ...
study	0	2	—	7.5 ...						germans	0	3	—	8.4 ...
document	0	2	—	8.3 ...						invader	0	3	—	8.6 ...
exhibition	0	3	—	9.4 ...						iraqi	0	3	—	8.6 ...

Q40. Word Sketch Difference: BNC利用
"art"と"arts"を使って単数名詞、複数名詞の使用傾向の違いを確認しなさい。同様に“fruit”と“fruits”の使用傾向も確認しなさい。

Q40.

BASIC **ADVANCED** ABOUT

compare ? 

Lemmas

Word forms 

Subcorpora

Lemma ?

art 

First word form ?

art 

Second word form ?

arts 

Part of speech ?

auto

noun 

verb

adjective

adverb

pronoun

Minimum frequency ?

auto

GO

Q41. Word Sketch Difference: BNC利用
Sketch Difference の Subcorpus機能を使って
動詞 "discover" が「fiction」と「science」でどのように使
われているか確認しなさい。

Q41.

Fictionには
Imaginative
を使用。

BASIC **ADVANCED** ABOUT

compare ? 

Lemmas
 Word forms
 Subcorpora

Lemma ?
discover

Subcorpus ?
Written Domain Imagi...  +

Subcorpus ?
Written Domain Infor...  +

Part of speech ?

auto
noun
verb
adjective
adverb
pronoun

Minimum frequency ?
auto

GO

Science には
Informative
を使用。

Q41.

Get more space

discover (Written Domain Imaginative) 1,980x | 6.0 | 4.0 | 2.0 | 0 | -2.0 | -4.0 | -6.0 | (Written Domain Informative) 7,959x

"discover" and/or ...			
come	—	4.6	...
use	—	6.4	...
identify	—	6.8	...
bring	—	7.3	...
learn	—	7.7	...
explore	—	7.7	...
develop	—	7.8	...
construct	—	7.9	...
appreciate	—	8.2	...
report	—	8.9	...
remedy	—	9.0	...
invent	—	9.4	...

objects of "discover"			
whereabouts	8.3	5.6	...
secret	8.0	7.2	...
identity	7.4	6.7	...
treasure	6.6	6.7	...
mistake	5.4	5.7	...
truth	7.3	8.1	...
existence	5.4	6.3	...
body	6.4	7.9	...
extent	5.3	6.9	...
talent	—	6.9	...
specie	—	7.0	...
cause	—	7.2	...

subjects of "discover"			
edouard	7.6	—	...
fabia	7.3	—	...
robbie	7.1	—	...
julia	6.9	—	...
father	3.9	3.3	...
mother	4.0	4.5	...
hahnemann	—	6.9	...
police	—	6.9	...
columbus	—	7.0	...
researcher	—	7.2	...
archaeologist	—	7.6	...
scientist	—	7.8	...

pronominal subjects of "discover"			
she	3.8	3.4	...
i	2.6	3.8	...
you	1.3	2.6	...
he	2.6	3.9	...
we	1.9	4.6	...
they	1.7	4.4	...
them	—	2.1	...
me	—	2.7	...
him	—	3.3	...
it	—	3.4	...
one	—	3.7	...

Q42. Word Sketch Difference: INACLE_WE利用 日本語話者の"make"の使用傾向を英語母語話者と比較しなさい。

Q42. Word Sketch Difference: INACLE_WE利用 日本語話者の"make"の使用傾向を英語母語話者と比較しなさい。

Word Sketch Differenceでは同一コーパス内のサブコーパス間で比較することができます。

The screenshot shows the 'WORD SKETCH DIFFERENCE' tool interface. At the top, there is a search bar containing 'ICNALE_WE'. Below it are three tabs: 'BASIC', 'ADVANCED', and 'ABOUT'. The 'ADVANCED' tab is selected. The interface is divided into several sections:

- compare?**: A section with three radio buttons: 'Lemmas', 'Word forms', and 'Subcorpora'. The 'Subcorpora' option is selected and circled in red.
- Lemma?**: A text input field containing the word 'make'.
- Subcorpus?**: Two subcorpus selection fields. The first is 'country_Japan' and the second is 'country_NativeSpeaker'. Both are circled in red.
- Part of speech?**: A dropdown menu with options: 'auto', 'adjective', 'adverb', 'noun', and 'verb'. 'auto' is selected.
- Minimum frequency?**: A text input field containing 'auto'.

At the bottom right, there is a red 'GO' button.

Q43. Word Sketch Difference: BNC利用 Sketch
Difference機能を使ってkeep + O Cとleave + O CのCの
部分にくる形容詞の違いを確認することによって、keepと
leaveの違いを確認しなさい。

Q43. Sketch Difference機能を使ってkeep + O Cとleave + O CのCの部分にくる形容詞の違いを確認することによって、keepとleaveの違いを確認しなさい。

確認してみると、keepとleaveの意味の違いが共起する語彙面から具体的に見えてきます。

leaveの後ろの補語にくる形容詞を見てみると、unから始まる形容詞など否定的な意味の言葉が多いことが分かります。



adjectives after "keep/leave" and noun				
warm	150	0	10.5	— ...
busy	114	0	10.1	— ...
happy	162	7	10.4	6.5 ...
alive	230	11	11.0	7.3 ...
clean	123	7	10.2	6.8 ...
intact	54	33	9.1	9.1 ...
open	213	250	10.6	11.4 ...
free	53	174	8.6	10.9 ...
untouched	0	29	—	9.0 ...
unchanged	0	30	—	9.1 ...
vulnerable	0	38	—	9.3 ...
late	0	49	—	9.7 ...

Q44. コーパスの大きさがThesaurusの精度を大きく左右することを確認するため "angry"を検索し、English Web 2020 と BNC それぞれの検索結果を比較しなさい。

Q44. コーパスの大きさがThesaurusの精度を大きく左右することを確認するため "angry"を検索し、English Web 2020 とBNCそれぞれの検索結果を比較しなさい

	Word	Frequency ?	Similarity ? ↓
1	furious	221,436	0.493 
2	mad	652,521	0.484 ...
3	unhappy	301,193	0.460 ...
4	anxious	396,886	0.410 ...
5	sad	1,152,385	0.400 ...
6	stupid	976,704	0.385 ...
7	happy	4,675,181	0.384 ...
8	crazy	1,118,913	0.383 ...
9	nervous	754,386	0.374 ...
10	desperate	582,661	0.372 ...

English Web 2020

Corpus size: 38,149,437,411

	Word	Frequency ?	Similarity ? ↓
1	unhappy	1,858	0.258 
2	anxious	2,952	0.245 ...
3	bitter	2,315	0.241 ...
4	sad	3,376	0.238 ...
5	tired	2,885	0.237 ...
6	furious	1,245	0.230 ...
7	nervous	2,876	0.214 ...
8	worried	742	0.195 ...
9	funny	4,318	0.192 ...
10	ill	3,368	0.192 ...

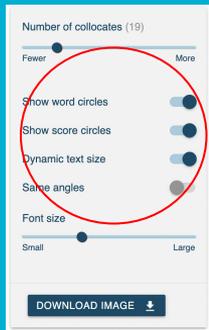
BNC

Corpus size: 96,134,547

Q45. 一般の類義語辞書に記載のない語でも Sketch EngineのThesaurus機能が使えることを確認するため English Web 2020 コーパスを使って "contumacious"を検索し、検索結果を Show visualization 機能で視覚化しなさい。

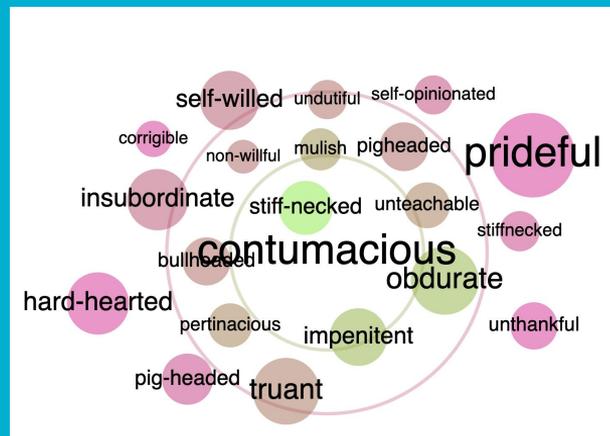
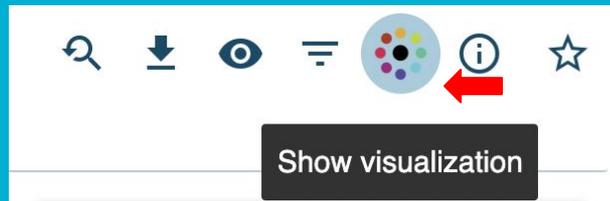
Q45. 一般の類義語辞書に記載のない語でもSketch EngineのThesaurus機能が使えることを確認するためEnglish Web 2020 コーパスを使って "contumacious"を検索し、検索結果をShow visualization機能で視覚化しなさい。

可視化された検索結果を見ると contumacious に最も近い類義語は stiff-necked であることが分かる。また、類義度は下がるが使用頻度が高い語が prideful であることがバブルの大きさから分かり、バブルのサイズが大きくなるほどより認知しやすい語に変化することを確認してください。



可視化された検索結果を好みに合わせて調整することが可能。

	Word	Frequency ?	Similarity ? ↓
1	stiff-necked	2,658	0.243 ...
2	obdurate	6,119	0.213 ...
3	impenitent	3,417	0.212 ...
4	mulish	875	0.193 ...
5	pertinacious	1,097	0.183 ...
6	unteachable	1,329	0.181 ...
7	truant	5,703	0.176 ...
8	bullheaded	1,643	0.170 ...
9	undutiful	584	0.170 ...
10	non-willful	191	0.168 ...



Q46. 名詞のdeclineの1つ前にくる形容詞の頻度リストをレマ単位で出し、1~5位までを確認しなさい。

Q46. 名詞のdeclineの1つ前にくる形容詞の頻度リストをレマ単位で出し、1~5位までを確認しなさい。

順番に、economic, sharp, rapid, steady, relativeが正解です。

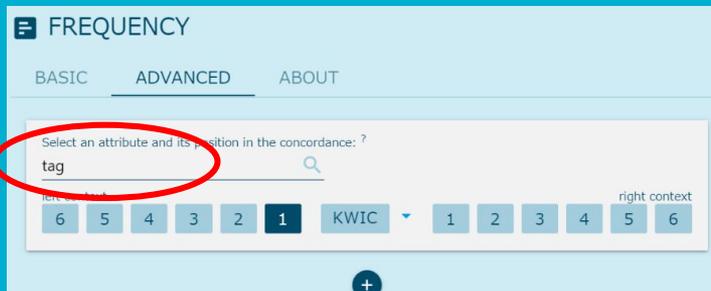
レマ検索で品詞を指定して、コンコーダンスを表示したら、FREQUENCYを選び、ADVANCEDタブでtagを選んだうえでleft contextの1をクリックします。

すると、declineの1つ左にくる品詞の頻度リストが表示されます。(DTは限定詞、JJは原級の形容詞を表しています。)

JJの右にあるボタンを選択し、Only thisをクリックすると、形容詞+declineのパターンが抽出されたコンコーダンスが表示されます。

※品詞タグについては

<https://www.sketchengine.eu/english-treetagger-pipeline-2/>



(39 items, 4,376 total frequency)

	Tag	Frequency ↓	Relative ?	
1	<input checked="" type="checkbox"/> DT	1,641	14.61	...
2	<input checked="" type="checkbox"/> JJ	1,376	12.25	Only this
3	<input type="checkbox"/> IN	397	3.53	Whole concordance but this
4	<input type="checkbox"/> NN	329	2.93	...
5	<input type="checkbox"/> CC	99	0.88	...
6	<input type="checkbox"/> PPZ	66	0.59	...

CONCORDANCE

British National Corpus (BNC)

lemma decline • 4,376
38.95 per million tokens • 0.0039%

filter [tag="JJ"] • 1,376
12.25 per million tokens • 0.0012%

Left context KWIC Right context

1	<input type="checkbox"/>	Written books a... then Australia... Then they began their mysterious and rapid decline... The last one was seen alive in 1979...
2	<input type="checkbox"/>	Written books a... to 50 companies may go out of business; 200 are expecting a serious decline in profitability, and many hundreds more are uncertain about their...
3	<input type="checkbox"/>	Written miscell... ness... With the USA, in the long term, set on irreversible decline the question is whether Japan or Europe can replace it as the wor...
4	<input type="checkbox"/>	Written miscell... , by the year 2,000, Europe should have reversed its relative economic decline and begun the investment in science and technology, in transport a...

Q47. 名詞のqualityの1つ前にくる形容詞(比較級・最上級を含む)の頻度リストを表記形で出し、1位～10位までを確かめなさい。

Q47. 名詞の quality の1つ前になる形容詞 (比較級・最上級を含む) の頻度リストを表記形で出し、1位～10位までを確かめなさい。

順番に、high, good, poor, better, top, highest, personal, higer, merchantable, bestが正解です。

前の問題では形容詞の中に比較級と最上級が入っていませんでした。そのため結果が少し変わっています。

前の問題と同じ要領で、品詞の頻度リストを出したら、JJだけでなくJJRとJJSにもチェックを入れコンコーダンスを表示します (JJRは形容詞の比較級、JJSは形容詞の最上級)。あとは先ほど同じ要領で頻度リストを表示しますが、その際 lemmaではなくwordを選びます。

CQLを使用して、もっと簡単に検索する方法がありますが、CQLの使用の仕方は後の問題で扱います。

(53 Items, 18,649 total frequency)

	Frequency ↓	Relative ↑
<input checked="" type="checkbox"/> JJ	5,312	47.28
<input type="checkbox"/> JJT	5,248	46.71
<input type="checkbox"/> IN	1,896	16.88
<input type="checkbox"/> NN	1,384	12.32
<input type="checkbox"/> CC	892	7.94
<input type="checkbox"/> PPZ	374	3.33
<input type="checkbox"/> JJR	349	3.11
<input type="checkbox"/> JJN	339	3.02
<input type="checkbox"/> NP	331	2.95
<input type="checkbox"/> VV	276	2.46
<input type="checkbox"/> JJV	267	2.38
<input checked="" type="checkbox"/> JJS	255	2.27
<input type="checkbox"/> SBN	241	2.15
<input type="checkbox"/> VVN	161	1.43
<input type="checkbox"/> CD	71	0.63
<input type="checkbox"/> :	62	0.55
<input type="checkbox"/> VVP	53	0.47
<input type="checkbox"/> NNZ	51	0.45

Show concordance for: **SELECTED (3)** | NOT SELECTED

FREQUENCY

BASIC ADVANCED ABOUT

Select an attribute and its position in the concordance: ?

word

left context: 6 5 4 3 2 1 KWIC 1 2 3 4 5 6 right context

Group by first column

GO

More presets

- KWIC WORD FORMS
- KWIC TAGS
- KWIC LEMMAS
- TEXT TYPES
- LINE DETAILS

Q48. CQL検索を利用し、"determine"を活用
形も含めて検索しなさい。

Q49. CQL検索を利用し、レマ単位でuniversity
とcollegeを同時に抽出しなさい。

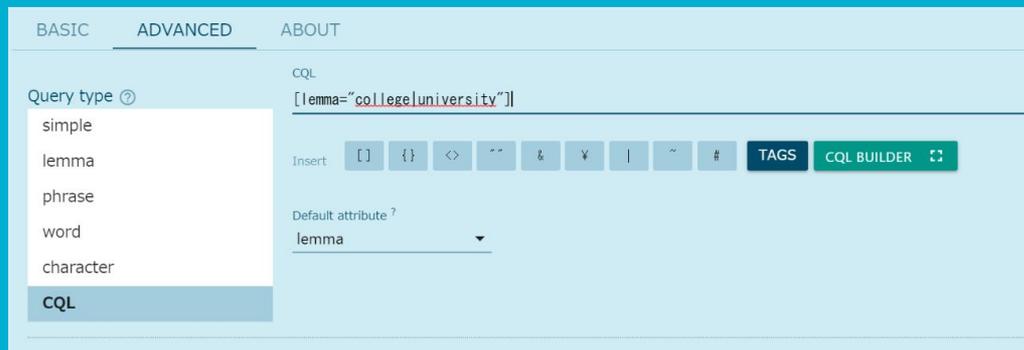
Q49. CQL検索を利用し、レマ単位でuniversityとcollegeを同時に抽出しなさい。

CQLでは| (パイプ, vertical bar) という記号が、orの意味を持ちます。

従って、

[lemma="university|college"]

とCQLを書くと、universityまたはcollegeを抽出するという指示になり、この2つの語を同時に検索することができます。



	Left context	KWIC	Right context
1	Written books about the history of molecular biophysics and Louise Johnson (Molecular Biophysics) were promoted to	university	readerships during Hilary Term
2	Written books about Professor Johnson, as she now is, was elected an Honorary Fellow of the	college	Jennifer Loach was awarded a British Academy / Lever
3	Written books about Olive Sayce, Fellow and Tutor in German since 1965, and	college	lecturer from 1946-65, retired at the end of the academic year
4	Written books about A noted authority on Medieval German, Olive was a most dedicated	college	tutor to generations of modern linguists, fostering and encouraging her pl
5	Written books about encouraging her pupils, and a fierce defender of modern languages in	college	and in the university She has been elected to an emeritus f
6	Written books about pupils, and a fierce defender of modern languages in college and in the	university	She has been elected to an emeritus fellowship, and has o
7	Written books about her over the gaudy weekend Almut Suerbaum, of the	university	of Munster, has joined us as the new Fellow in German; as a specialist i
8	Written books about Proctors' tasks Joanna's experience will be invaluable to the	college	when she relinquishes office at the end of the Proctorial year

Q50. CQL検索を使用し、be動詞とhaveを除いた動詞の原形のみを抽出しなさい。

Q50. CQL検索を使用し、be動詞とhaveを除いた動詞の原形のみを抽出しなさい。

[tag="VV"]というCQLを書きます。

品詞タグは<https://www.sketchengine.eu/english-treetagger-pipeline-2/>で確認できます。

動詞を表すタグはVから始まります。VBはbe動詞、VHはhave、VVはそれ以外の動詞を表します。VB、VH、VVだけだとそれぞれの原形(base form)を表し、その後ろにさらにD、G、Nなどを続けると様々な活用形を指定できます。Dは過去形、Gはing形、Nは過去分詞形、Pは現在形(3単元を除く)、Zは3単現を表します、

従って、例えばVBDIはbe動詞の過去形(was, were)、VHNIはhaveの過去分詞形(had)、VVGは一般動詞のing形(taking, askingなど)を表すこととなります。

VB	verb be, base form	be
VBD	verb be, past tense	was, were
VBG	verb be, gerund/present participle	being
VBN	verb be, past participle	been
VBP	verb be, present, non-3d person	am, are
VBZ	verb be, 3rd person sing. present	is
VH	verb have, base form	have
VHD	verb have, past tense	had
VHG	verb have, gerund/present participle	having
VHN	verb have, past participle	had
VHP	verb have, sing. present, non-3d	have
VHZ	verb have, 3rd person sing. present	has
VV	verb, base form	take
VD	verb, past tense	took
VVG	verb, gerund/present participle	taking
VVN	verb, past participle	taken
VVP	verb, present, not 3rd person	take
VVZ	verb, 3rd person sing. present	takes

Q51. CQL検索を利用し、be動詞とhaveを除いた動詞を、その活用形も含めて抽出しなさい。

Q51. CQL検索を利用し、be動詞とhaveを除いた動詞を、その活用形も含めて抽出しなさい。

[tag="VV.?"]というCQLを書きます。

CQLの""の中では正規表現を使うことができます。

「.」(ドット)は任意の一文字を表します。「?」は直前の文字が0個か1個という意味です。

従って、VV.?はVVの後に何も来ない、もしくは何か1文字くるとことを示します。つまりこの書き方で、VV, VVD, VVG, VVN, VVP, VVZを動詞に指定できるので、be動詞とhave以外の全ての動詞の全ての活用形を同時に抽出できます。

このように品詞タグに正規表現を使うことによって、うまく検索したい品詞とその活用形を指定することができます。

Q52. CQL検索を利用し、普通名詞の単数形と複数形を抽出しなさい。その際、固有名詞を、所有格の名詞を除けるように工夫したCQLを書きなさい。

Q52. CQL検索を利用し、普通名詞の単数形と複数形を抽出しなさい。その際、固有名詞を、所有格の名詞を除けるように工夫したCQLを書きなさい。

[tag="NNS?"]というCQLを書きます。

よく使うタグはCQL入力欄の右側になるTAGSの中から選んで、コピーすることができます。

ここにnounのタグとして、N.*とありますが、これはNから始まるタグを全て含むので固有名詞や所有格の名詞も含まれます。

単数形の普通名詞のタグはNN、複数形の普通名詞のタグはNNSなのでこの二つのみを同時に指定するには上記のように書くことができます。正規表現で「?」は一つ前の文字が0~1個という意味なので、Sがあってもなくてもよいということになり、NNとNNSを同時に指定することになります。

ちなみに、[tag="NN|NNS"]と書いても同じことです。



COMMON TAGS

adjective	J.*	
adverb	RB.?	
conjunction	CC	
determiner	DT	
noun	N.*	
numeral	CD	
particle	RP	
preposition	IN	
pronoun	PP.?	
verb	V.*	

view help for [ALL TAGS](#)

CLOSE

Q53. CQL検索を利用し、単語を指定せず、動詞＋名詞のパターンを抽出しなさい。

Q53. CQL検索を利用し、単語を指定せず、動詞＋名詞のパターンを抽出しなさい。

[tag="品詞タグ"]という式で、品詞を指定した検索ができます。

[tag="V.*"]は全ての動詞を指定、[tag="N.*"]は全ての名詞を指定します。

[tag="V.*"][tag="N.*"]と続けて打ち込めば、動詞＋名詞のパターンを抽出できます。

FREQUENCYでレマでKWICの頻度を確認すると、BNCではsee pageというフレーズが最も多いことが確認できます。

CONCORDANCE British National Corpus (BNC) Get mo

BASIC ADVANCED ABOUT

Query type ①

simple lemma phrase word character CQL

CQL [tag="V.*"][tag="N.*"]

Insert [] { } < > * & ¥ | ~ # TAGS

CQL BUILDER ⇄

Default attribute ? lemma

Subcorpus ① none (the whole corpus) + Macro ? none

Filter context ① ▾

Text types ? ▾

GO

CQL 1: complex cor... an introduction to corpus linguistics language SKETCH ENGINE www.sketchengine.eu CQL manual

Q54. CQL検索を利用し、副詞のwayを抽出し、右1語でアルファベット順にソートしなさい。

Q54. CQL検索を利用し、副詞のwayを抽出し、右1語でアルファベット順にソートしなさい。

&を使えば、複数の条件を同時に指定できます
[lemma="way"]と検索すると、品詞に関係なく全てのwayが抽出されるので、
[lemma="way"&tag="RB.?"]とすることで、副詞のwayのみを抽出することができます。

ソート機能を使えば、指定した位置の語を基準にして、アルファベット順にコンコーダンスラインを並べ替えられます。

wayの右1語でソートすれば、どのような語が多くwayの右に来るかわかりやすくなります。



CONCORDANCE British National Corpus (BNC)

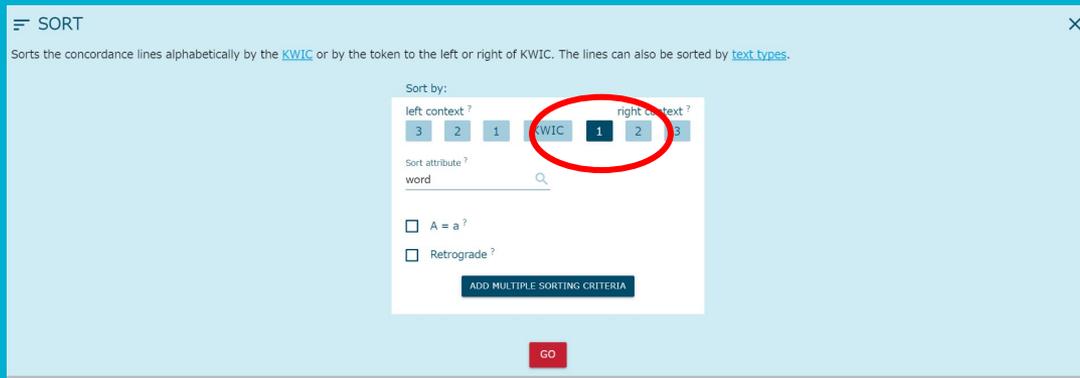
CQL [lemma="way"&tag="RB.?"] 301 | Sort word, word x

Left context KWIC Right context

1 Spoken demograp... My handbag 's through that </s><s> Turn the tape off . </s><s> way , it 'll be quicker </s><s> Are you turning the tape off ? </s><s> O

2 Written books a... iccustomed to the dark , and she could see , as she peered down it either way , that to the left this ro... opened out . </s><s> A square , perhap

3 Spoken demograp... round and stuff </s><s> fucking hilarious </s><s> yeah </s><s> they way , way he I mean he 's took a bite out of it or something </s><s> he



SORT

Sorts the concordance lines alphabetically by the KWIC or by the token to the left or right of KWIC. The lines can also be sorted by [text types](#).

Sort by:

left context ? 3 2 1 KWIC 1 2 3 right context ?

Sort attribute ?

word

A = a ?

Retrograde ?

ADD MULTIPLE SORTING CRITERIA

GO

Q55. CQL検索で、動詞のcharge + 名詞の
パターンを抽出しなさい。

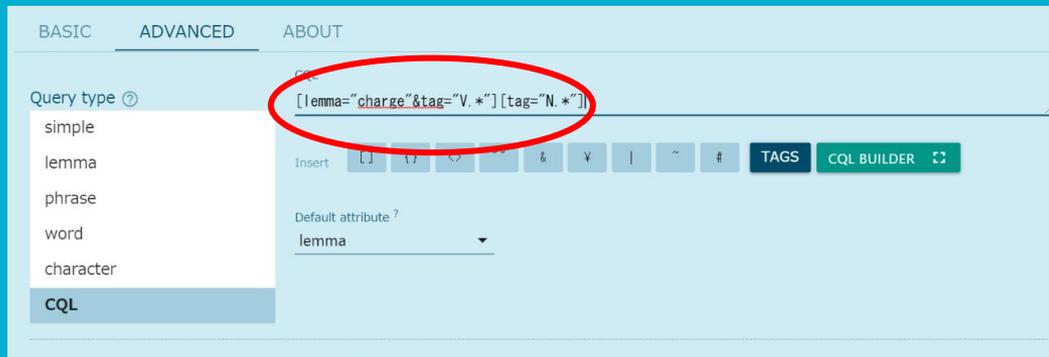
Q55. CQL検索で、動詞のcharge+ 名詞のパターンを抽出しなさい。

&を使えば、複数の条件を同時に指定できます
[lemma="charge"]と検索すると、動詞chargeも
名詞chargeも抽出されるので、動詞のcauseを
指定する場合は、[lemma="charge" &
tag="V.*"] と検索します。

```
[lemma="charge"&tag="V.*"][tag="N.*"]
```

と検索すれば、動詞charge + 名詞のパターンを
抽出することができます。

FREQUENCYでレマ単位でKWICの頻度を確認
するとcharge particle, charge fee, charge
interstというフレーズが多いことが分かります。



Q56. CQL検索で名詞のsenseから始まる3単語連鎖を検索しなさい。

Q56. CQL検索で名詞のsenseから始まる3単語連鎖を検索しなさい。

[lemma="sense"&tag="N.*"][] [2]

というCQLを書きます。

[]の中に何も書かないと全ての語を指定することになります[]の後ろは{2}は直前のものが2回続くという意味になります。従って[] {2}は「何でもいいので2語」という意味になります。

従って上記のCQLは「活用形を含む名詞のsenseの後ろに何か2語が続く」という意味になり、名詞のsenseから始まる3語のフレーズ(3単語連鎖)を指定したことになります。

FREQUENCYでKWICの頻度をレマ単位で出すと、sense of the, sense of humour, sense in which, sense of aなどが上位に来ており、これらのフレーズがよく使われることが分かります。

CONCORDANCE British National Corpus (BNC)

BASIC **ADVANCED** ABOUT

Query type ?

- simple
- lemma
- phrase
- word
- character
- CQL**

CQL [lemma="sense"&tag="N.*"][] [2]

Insert [] {} <> "" & ¥ | ~ # TAGS

CQL BUILDER

Default attribute ? lemma

**Q.57. CQL検索で名詞のcapacityから始まる
3~5単語連鎖を検索しなさい。**

Q57. CQL検索で名詞のcapacityから始まる3~5単語連鎖を検索しなさい。

[lemma="capacity"&tag="N.*"]{2,4}

というCQLを書きます。

{2,4}は直前の方が2~4語続くという意味になります。従って、[[{2,4}]は「何でもいから2~4単語」という意味になります。

従って一番上のCQLは「活用形を含む名詞のcapacityの後ろに何か2~4語続く」という意味になり、名詞のcapacityから始まる3~5語のフレーズ(3~5単語連鎖)を指定したことになります。

コンコーダンスを見ると分かりますが、これだと文を跨いだものまで含まれています。

1	<input type="checkbox"/>	Written books a... satory increases in natural gas and nuclear generating	capacity ; and	the planting of large areas of forest , which would act a
2	<input type="checkbox"/>	Written books a... satory increases in natural gas and nuclear generating	capacity ; and the	planting of large areas of forest , which would act as a
3	<input type="checkbox"/>	Written books a... satory increases in natural gas and nuclear generating	capacity ; and the planting	of large areas of forest , which would act as a " sink
4	<input type="checkbox"/>	Written books a... ar power's share of the country's electricity generating	capacity to 17	per cent . </s><s> A public opinion poll published
5	<input type="checkbox"/>	Written books a... ar power's share of the country's electricity generating	capacity to 17 per	cent </s><s> A public opinion poll published at t
6	<input type="checkbox"/>	Written books a... ar power's share of the country's electricity generating	capacity to 17 per cent	</s><s> A public opinion poll published at the e
7	<input type="checkbox"/>	Written miscell... : group's malt distilleries . </s><s> The combined	capacity of the	maltings is 210,000 tonnes , making United Distillers
8	<input type="checkbox"/>	Written miscell... : group's malt distilleries . </s><s> The combined	capacity of the maltings	is 210,000 tonnes , making United Distillers the secur
9	<input type="checkbox"/>	Written miscell... : group's malt distilleries . </s><s> The combined	capacity of the maltings is	210,000 tonnes , making United Distillers the second la
10	<input type="checkbox"/>	Written miscell... : illeries producing almost a third of the total industry	capacity . </s><s> United	Distillers produces malt whiskies in all four whisky-pro
11	<input type="checkbox"/>	Written miscell... : illeries producing almost a third of the total industry	capacity . </s><s> United Distillers	produces malt whiskies in all four whisky-producing ' r
12	<input type="checkbox"/>	Written miscell... : illeries producing almost a third of the total industry	capacity . </s><s> United Distillers produces	malt whiskies in all four whisky-producing ' regions ' .
13	<input type="checkbox"/>	Written miscell... duate admissions and tutorial teaching in their college	capacities . </s><s> Graduate	students must also secure a college place , but each le
14	<input type="checkbox"/>	Written miscell... duate admissions and tutorial teaching in their college	capacities . </s><s> Graduate students	must also secure a college place , but each laboratory
15	<input type="checkbox"/>	Written miscell... duate admissions and tutorial teaching in their college	capacities . </s><s> Graduate students must	also secure a college place , but each laboratory and t
16	<input type="checkbox"/>	Written miscell... demand for Guinness began to outstrip the production	capacities of St.	James's Gate . </s><s> So the first overseas brew
17	<input type="checkbox"/>	Written miscell... demand for Guinness began to outstrip the production	capacities of St. James's	Gate . </s><s> So the first overseas brewery was l
18	<input type="checkbox"/>	Written miscell... demand for Guinness began to outstrip the production	capacities of St. James's Gate	. </s><s> So the first overseas brewery was built c
19	<input type="checkbox"/>	Written books a... Nuclear Power </s><s> USSR plans to triple nuclear	capacity by 2000	</s><s> According to Soviet Minister for Atomic Energ
20	<input type="checkbox"/>	Written books a... Nuclear Power </s><s> USSR plans to triple nuclear	capacity by 2000 </s><s> According	to Soviet Minister for Atomic Energy , Vitaly Konopalov
21	<input type="checkbox"/>	Written books a... Nuclear Power </s><s> USSR plans to triple nuclear	capacity by 2000 </s><s> According to	Soviet Minister for Atomic Energy , Vitaly Konopalov , t
22	<input type="checkbox"/>	Written books a... erment is planning a dramatic increase in nuclear	capacity and the	sale of uranium on world markets </s><s> Quoty

Q58. 動詞のenjoyの目的語になる名詞の頻度表を作成しなさい。

Q58. 動動詞のenjoyの目的語になる名詞の頻度表を作成しなさい。

[lemma="enjoy"&tag="V.*"][tag="DT|J.*|[CNP].*Z"]{0,2}{tag="N.*"]

というCQLで検索します。[CNP].*Zは所有格を指定しています

頻度表を作成するにはFREQUENCYの機能を使用します。活用形もひとまとめにして出すためにattributeはlemmaを選択し、頻度を出す語の場所をLAST KWIC WORDに設定します。

LAST KWIC WORDを指定すると、KWIC(検索した文字列)の中の最後の単語を指定して、頻度表を出すことができます。今回はCQLの中で[tag="N.*"]が最後にあり、これが基本的にはenjoyの目的語になる名詞なので、この指定の仕方ではenjoyの目的語になる頻度表を出すことができます。

The screenshot shows the 'FREQUENCY' tool interface. The search query is entered in the search bar. The 'attribute' dropdown is set to 'lemma' and the 'position' dropdown is set to 'LAST KWIC WORD'. The 'GO' button is visible at the bottom.

(1,943 items, 7,283 total frequency)

	Lemma	Frequency ↓	Relative ?	
1	<input type="checkbox"/> life	196	1.74	...
2	<input type="checkbox"/> company	118	1.05	...
3	<input type="checkbox"/> benefit	92	0.82	...
4	<input type="checkbox"/> success	82	0.73	...
5	<input type="checkbox"/> view	80	0.71	...
6	<input type="checkbox"/> freedom	75	0.67	...
7	<input type="checkbox"/> support	75	0.67	...
8	<input type="checkbox"/> work	74	0.66	...
9	<input type="checkbox"/> day	73	0.65	...
10	<input type="checkbox"/> experience	68	0.61	...
11	<input type="checkbox"/> holiday	61	0.54	...
12	<input type="checkbox"/> evening	53	0.47	...
13	<input type="checkbox"/> game	51	0.45	...

Q59. BNCを利用し、get passive(get + 過去分詞)の使用頻度を話し言葉と書き言葉で比較しなさい。

Q59. BNCを利用し、get passive(get + 過去分詞)の使用頻度を書き言葉と話し言葉で比較しなさい。

まずは、get + 過去分詞を抽出するために、

`[lemma="get"][tag="VVN"]`

というCQLを書きます。(VVNは動詞の過去分詞形を表すタグです。)

書き言葉と話し言葉での頻度をそれぞれ出すためにはサブコーパス指定をします。

SkE上ではBNCのサブコーパスがプリセットで入っています。SubcorpusのSpoken transcriptsを選択すれば、話し言葉を指定して検索でき、Written textsを指定すれば書き言葉を指定して検索します。

コンコーダンスの左上に、検索したQueryのヒット数と、100万語あたりの頻度(per million tokens)が表示されています。

CONCORDANCE British National Corpus (BNC)

BASIC ADVANCED ABOUT

Query type ①

- simple
- lemma
- phrase
- word
- character
- CQL

CQL [lemma="get"][tag="VVN"]

Insert [] { } < > * & % | ~ # TAGS

CQL BUILDER ⇄

Default attribute ? lemma

Subcorpus ② none (the whole corpus) 🔒

Macro ? none 🔍 🔒 ⚙

Filter context ③ ⌵

Text types ? ⌵

GO

CONCORDANCE British National Corpus (BNC)

Spoken transcripts > CQL [lemma="get"][tag="VVN"] • 3,207
271.79 per million tokens • 0.0029%

Details Left context KWIC Right context

1	Spoken context-... s><s> Not going to sing . </s><s> I promise not , well I might	get carried away .
2	Spoken context-... ome other business and I felt I got off the hook here but in fact that	got dealt with very q
3	Spoken context-... ecause he 's full of aggression is n't he . </s><s> Well he does	get killed off in that c
4	Spoken context-... e time in the future , you do have er , some bad health , or perhaps	get involved in an accid

Q60. コンコーダンスで3文字から成る形容詞
を検索しなさい。

Q60. コンコーダンスで3文字から成る形容詞を検索しなさい。

`[word="..."&tag="JJ"]`

というCQLを書きます。

LEMMA, PHRASE, WORD, CQL検索では正規表現 (regular expressions)を使用することができます。これによりさらに様々な検索を行うことができます。

正規表現では「.(ドット)は任意の1文字を表します。従って「...」は、3文字の単語を指定することになります。

CQL検索では、「」の中に正規表現を使用することができます。上記の検索式の `word="..."` の部分は3文字の単語を指定しています。従って上記の検索式は3文字の形容詞を指定することになります。(JJは形容詞のタグ)

Q61. succeedとその派生形をまとめて検索しなさい。(successful, success, successive, successionなど)

Q61. succeedとその派生形をまとめて検索しなさい。(successful, success, successive, succession など)

[word="succe.*"] とCQLを書きます。

正規表現では「*」(アスタリスク)は「直前の文字を0回以上繰り返す」という意味です。

「.*」は字数を指定せずに全ての文字列を指定することになります。succe.*とすれば字数を指定せずに succeから始まる全ての単語を検索することができ、succeedの派生語をまとめて抽出できます。(ただしunsuccessfulなどは含まれない。

他にも、例えば、dis.* とすれば、disから始まる全ての単語を抽出できます。

Q62. 語尾が-tionもしくは-sionになる名詞を
その複数形もふくめて検索しなさい。

Q62. -tionもしくは-sionで終わる名詞をその複数形も含めて検索しなさい。

[lemma="*.tion|.sion"&tag="N.*"]

とCQLを書きます。

*.tionとすれば、語尾がtionになる全ての単語を指定できます。

複数形も含めるために、lemmaで検索し、tag="N.*"で名詞を指定します。

CONCORDANCE British National Corpus (BNC)

CQL [lemma="*.tion|.sion"&tag="N.*"] 1,570,718
13,981.11 per million tokens • 1.4%

	Left context	KWIC	Right context
1	Written books a... l (1979) and Brilliant (1974) .	recommendations	of appropriate taste or choice of style are few and very general .
2	Written books a... rned with the architectural aspects of mosaic , e.g. the technicalities of	foundations	Pliny (Hist . 30.60.64) , meanw
3	Written books a... only slight differences , e.g. the types of pavements to which particular	foundations	are applicable . Even commentaries on those crafts which
4	Written books a... mosaic , such as Lucian's Dream , are of little assistance .	Inscriptions	, within mosaics and elsewhere , provide a richer source of informati
5	Written books a... nscriptions , within mosaics and elsewhere , provide a richer source of	information	. These do require a certain degree of inference , but it
6	Written books a... uire a certain degree of inference , but it is clear that under some	conditions	such remains can suggest the name of a sponsor , the reason for spon:
7	Written books a... s of individual craftsmen are attested (although , as is shown by the	inscription	from Timagd , even differing ideas of the practice in which the craftsma
8	Written books a... , the potters' stamps of second century , Gallic , Samian ware : for the	organisation	of style , consciousness of individual or specific style , or written marks
9	Written books a... slight and difficult to interpret .	inscriptions	are not very helpful . Most were intended to be exulting c
10	Written books a... </s><s> Most were intended to be exulting or to act as aids to the	recognition	of individual figures . Only at Bignor , Sussex (no. 5 , A)
11	Written books a... uld have been more explicit (pl. 37) .	consideration	of the archaeological (as opposed to the specifically epigraphic) evide
12	Written books a... > B - indicate little more than the immediate place of work during the	construction	of the pavements of a single building . In most instances
13	Written books a... </s><s> 2.2 </s><s> It appears that only limited insights into the	organisation	of style are to be gained from those sources - commentaries , inscripti
14	Written books a... nisation of style are to be gained from those sources - commentaries ,	inscriptions	, and socio-economic indicators - which reflect upon style .
15	Written books a... economic indicators - which reflect upon style .	limitation	therefore , one has to rely on an interpretation of the relationships bet
16	Written books a... </s><s> Given this limitation therefore , one has to rely on an	interpretation	of the relationships between the mosaics themselves .
17	Written books a... relationships between the mosaics themselves .	interpretation	involves us , immediately , with a consideration of style and , by implic
18	Written books a... </s><s> Such an interpretation involves us , immediately , with a	consideration	of style and , by implication , " form " (that which has been included

Q63. BNCで、multinationalという単語を含む文の数と段落の数を調べなさい。

Q63. BNCで、multinationalという単語を含む文の数と段落の数を調べなさい。

<s/>containing[word="multinational"] というCQLと、

<p/>containing[word="multinational"] というCQLを書きます。

containingというCQLは「右側のものを含む」という意味です。1つめのCQLではmultinationalという単語を含むセンテンス、2つめはmultinationalという単語を含むパラグラフを指定しています。

CONCORDANCE British National Corpus (BNC)

CQL <s/>containing[word="multinational"] • 544
4.84 per million tokens • 0.00048%

Details sentence

- Written books a... <s> Inhutani II has been brought in to replace the multinational Scott company , which pulled out of the project at the end of 1989 after an intensive campaign by environmental groups , centred around the Women's Environmental Network and Survival International . </s>
- Written books a... <s> The trees were developed by Spanish engineer Antonio Ibanez , who claims to have interested an Italian multinational plus the French and more than one Arab government in his invention . </p></s>
- Written miscell... <s><p> Large multinational corporations produce and acquire vast volumes of information in the course of their business . </s>
- Written miscell... <s><p> Large multinational corporations produce and acquire vast volumes of information in the course of their business . </s>
- Written miscell... <s> Such databases are being used in a wide range of contexts ; for example , as an international library service (eg . Reuters textline and Butterworth's Lexis) ; or as a store of specialised technical information within multinational companies , industry groupings or specialised customer / supplier relationships (eg . the CALS initiative of the US Department of Defence) . </s>
- Written books a... <s><p> Communists will campaign for democratic control of multinational companies , for safety at work and job security . </s>
- Written books a... <s><p> One of the architects of the bill , Dr Gordon Guymer , head of the Queensland Herbarium , said it would ensure that " when multinational companies come here to screen the biota for new drugs , the state will have some ownership over that resource " . </s>

Q64. bigを含む形容詞が2~3個つく名詞句を抽出しなさい。

Q64. bigを含む形容詞が2~3個つく名詞句を抽出しなさい。

`[tag="J.*"]{2,3}[tag="N.*"]containing[lemma="big"&tag="J.*"]`

というCQLを書きます。

containingの前には、</s/>や<p/>といったstructureを表すものだけでなく、CQLを書くことができます。

`[tag="J.*"]{2,3}[tag="N.*"]`の部分は、2~3つ形容詞がつく名詞を示しており、これに`containing[lemma="big"&tag="J.*"]`を続けることで、bigを含む形容詞が2~3個つく名詞句を抽出することができます。2つ以上形容詞が付く場合は、当該の形容詞が何番目につくか分からないので、このように検索することで、もれなく検索したい形容詞を含んだ名詞句を抽出することができます。

CONCORDANCE British National Corpus (BNC)

CQL `[tag="J.*"]{2,3}[tag="N.*"]containing[lemma="big"&tag="J.*"]` 3,654
32.52 per million tokens • 0.0033%

	Left context	KWIC	Right context
1	Written miscell... /s><p> The School of Geography at Oxford is one of the	biggest autonomous Geography	departments in any British university, with two professors
2	Written miscell... rited from the 17th century Huguenot brewers) are essentially	great big strainers	or filtering vessels where the mash is allowed to 'stand-in',
3	Written books a... tributor to the greenhouse effect . </p></s><p><p> The	biggest single curb	on deforestation has been the government's temporary seizure
4	Written books a... hat contain its hamburgers . </s><s> For years, the US's	biggest fast food	company has staunchly defended its right to use the plastic bo
5	Written books a... s><p> The debate over the salmon is likely to prove an even	bigger political issue	than that of the future of the northern spotted owl (see ED No.
6	Spoken context... ould be really nice and you watch it outdoors and it 's in this	big stately home	which is in it 's own grounds and there 's gift shops and rest
7	Spoken context... the bars but you know there 's lo it 's beautiful and like a	big stately home	you can wander round the gardens for a bit and then go and we
8	Spoken context... u walked around to look at the flat, would you find a lake and	whacking big golf	course in the middle of it ? </s><s> no </s><s> now let's
9	Written books a... /s><s><p> As Ken recalled : ' Loot was an appalling flop , a	great big flop </s><s> It was like a ship that set out with the navigator
10	Spoken context... ey did n't have enough people with them, they did n't have a	big enough army	</s><s> So they decided to retreat . </s><s> What
11	Spoken context... h I know . </s><s> Say yeah you can make you know the	big bright lights	and things like that . </s><s> Which again if you accept </
12	Spoken context... imagine how much they 're cutting . </s><s> The also the	big rotary blades	on 'em are throwing all the grass over the pavement and all o
13	Spoken context... with a proposal for Road but it will have to be compared with	other big schemes	so the scheme we are proposing now is the one that 's befor
14	Spoken context... ie was an intellectual in his way, you know I mean he was a	big bright cookie	and but his was in terms of word count because he had a use
15	Spoken context... onday working lunch . </s><s> Is there any chance of a	bigger airier room	, or is it just </s><s> Whenever possible I try to get a L
16	Spoken context... er, last year Alcuin J C R booked Kingmaker as part of their	massive big mega-media	global world tour . </s><s> And then a fortnight before th
17	Spoken context... ed is probably big enough, twenty five might be is more or	less big enough	</s><s> Mm . </s><s> Er five, say we can have five
18	Spoken context... are right of four hundred and ten now you do n't need to be a	real big mathematician	, even at this time of night, but you, if you have the two hund

Q65. 文頭にimagineが来る文を抽出しなさい。

Q65. 文頭にimagineが来る文を抽出しなさい。

<s>[lemma="imagine"]

というCQLを書きます。

<s>はセンテンスを表す記号です。Sketch Engineに挙げられているデータは<s></s>というタグで囲むことによってその範囲が一つのセンテンスであることを示しています。

CQLの最初に<s>を置くことによって、文頭であることを指定することができます。

CONCORDANCE British National Corpus (BNC)

CQL <s>[lemma="imagine"] 618
5.5 per million tokens 0.00055%

Details sentence

1 Written miscell... <s> **Imagine** your favourite operatic tenor is singing your favourite aria just for you, perfect bel canto, every syllable clear, yet hushed and honeyed, subduing a richly passionate voice and modulating it with an effect of caged power ! </s>

2 Written miscell... <s> **IMAGINE** HOW A CELT FELT WITHOUT GUINNESS </s>

3 Written miscell... <s> **Imagine** what might be done with the collected works of Alexander Hamilton when the Founding Fathers project eventually provides them on CD-ROM sometime in the not too distant future . </s>

4 Spoken context-... <s> **Imagine** having to lift that from a </s>

5 Spoken context-... <s> **Imagine** with me Chairman if you will the thoughts of someone trapped by some misfortune in the midst of the wreckage of a road traffic accident on the M 1 motorway Southbound between junctions twenty one and twenty that 's at Nutterworth or Northbound between twenty one and twenty two or Westbound on the M 69 they 'll be thinking the fire brigade 'll be here in a minute or two . </s>

6 Written books a... <s> **Imagine** a similar hilltop in the lovely countryside of Aviemore and a small stone-built lodge providing shelter for a corporal and two airmen at one of the quadrants of a bombing range that Kinloss used for practice . </s>

7 Written books a... <s> **Imagine** that a series of pebbles is being dropped in quick succession into the middle of a pond, so that waves are continuously radiating out from the middle . </s>

8 Written books a... <s> **Imagine**, first, a horseshoe bat giving out a continuous hum of ultrasound as it flies fast towards a still object, like a tree . </s>

9 Written books a... <s> **Imagine** the " bee-flower " of Figure 5 in colour . </s>

10 Spoken context-... <s> **Imagine** if you were a vixen scared and frightened, seeing her cubs destroyed . </s>

11 Written books a... <s><p> **Imagine** a pragmatist judge deciding McLoughlin . </s>

12 Spoken context-... <s> **Imagine** yourself in a night club . </s>

13 Spoken context-... <s> **imagines** a forest you think and of course there are real forests out there somewhere are n't there ? </s>

Q66. Concordanceのコロケーションを機能を使って、形容詞harshの1語後ろにくる名詞のリストをTスコアでソートして出さない。

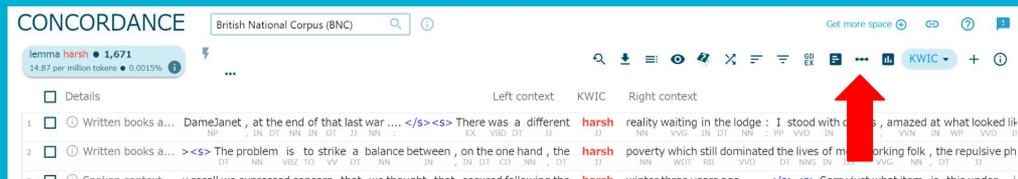
Q66. Concordanceのコロケーション機能を使って、形容詞 harshの1語後ろにくる名詞のリストをTスコアでソートして出さない。

Concordanceのコロケーション機能を使って、コロケーションを調べることができます。

検索語の1語後ろに来る語を調べたい場合は範囲を図のように、右の1に指定します。

コンコーダンスのコロケーション機能では、様々な品詞が混ざってしまうので、harshの1語右にくる名詞を指定したい場合は、フィルター機能を使って、harshの1語後ろに名詞がくるもののみを抽出したうえでコロケーション機能を使います。

コロケーション機能では、単なる共起頻度だけではなく、T-scoreやMIやLogDiceなどといった共起強度を測定する統計指標を利用することができます。T-scoreは高頻度語を評価しやすい、MIは低頻度だが特定の語との結びつきが強いコロケーションを評価しやすいなど、指標ごとに特徴があります。



CONCORDANCE British National Corpus (BNC)

lemma harsh • 1,671
14.87 per million tokens • 0.0015%

Get more space

Left context KWIC Right context

1 Written books a... DameJanet, at the end of that last war ... </s><s> There was a different harsh reality waiting in the lodge: I stood with c... , amazed at what looked li...

2 Written books a... ><s> The problem is to strike a balance between, on the one hand, the harsh poverty which still dominated the lives of m... rking folk, the repulsive ph...

3 Spoken context... u recall we expressed concern, that we thought that secured following the harsh winter three years ago. /><s> Sorry just what tem... is this under...



COLLOCATIONS

BASIC ADVANCED ABOUT

Attribute? Range?

word -5 -4 -3 -2 1 KWIC 1 2 3 4 5

GO

Collocations [CHANGE CRITERIA](#) [BACK TO CONCORDANCE](#)

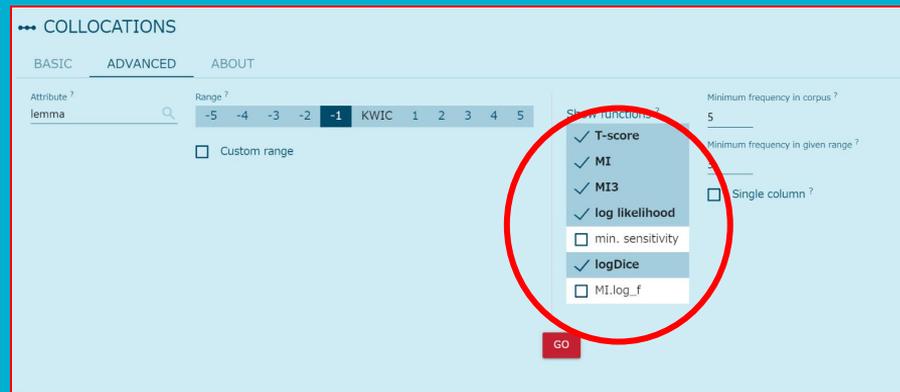
	Lemma	Cooccurrences?	Candidates?	T-score↓	MI	LogDice	
1	<input type="checkbox"/>	reality	74	7,135	8.60	10.24	8.23 ...
2	<input type="checkbox"/>	word	60	41,830	7.70	7.38	5.52 ...
3	<input type="checkbox"/>	voice	28	19,347	5.26	7.40	5.50 ...
4	<input type="checkbox"/>	treatment	26	12,738	5.08	7.89	5.96 ...
5	<input type="checkbox"/>	laugh	21	10,642	4.56	7.84	5.89 ...
6	<input type="checkbox"/>	environment	20	11,376	4.45	7.68	5.73 ...

Q67. Concordanceのコロケーション機能を使って、名詞のurgeの1語前にくる語をTスコアとMIでソートし、どのような語が上位に来るか比べなさい。

Q67. Concordanceのコロケーション機能を使って、名詞のurgeの1語前にくる語をTスコアとMIでソートし、どのような語が上位に来るか比べなさい。

コンコーダンスのコロケーション機能では同時に複数のコロケーション統計を選択することができます。

各指標の特徴をつかむためには、各指標でコロケーションを抽出するとどのような語が上位に現れるのかを実際に見てみる重要です。



	Lemma	Cooccurrences ?	Candidates ?	T-score	MI ↓	MI3	Log likelihood	LogDice	
1	<input type="checkbox"/> uncontrollable	9	212	3.00	12.63	18.97	140.04	8.25	...
2	<input type="checkbox"/> instinctive	6	479	2.45	10.87	16.04	78.51	7.32	...
3	<input type="checkbox"/> irresistible	6	483	2.45	10.85	16.02	78.41	7.31	...
4	<input type="checkbox"/> overwhelming	14	1,275	3.74	10.68	18.29	179.63	7.82	...
5	<input type="checkbox"/> addictive	4	558	2.00	10.06	14.06	47.85	6.64	...
6	<input type="checkbox"/> biological	6	1,871	2.44	8.90	14.07	62.12	6.23	...
7	<input type="checkbox"/> sexual	18	6,663	4.23	8.65	16.99	180.49	6.31	...
8	<input type="checkbox"/> creative	6	2,387	2.44	8.55	13.72	59.20	5.97	...
9	<input type="checkbox"/> sudden	10	4,099	3.15	8.51	15.15	98.13	6.08	...

Q68. 動詞のdriveの右2～3語の位置に来る
形容詞の頻度を確かめなさい。

Q68. 動詞のdriveの右2～3語の位置に来る形容詞の頻度を確かめなさい。

FILTER機能を使えば、複雑なCQLを書かずとも、見たい表現をうまく抽出することができます。

動詞のdriveの2～3語の位置にくる形容詞を調べたい場合、Keep linesをcontainingにし、Query type: CQLにし、形容詞を指定するCQLを書きます。その後、下のRangeで2・3を選択肢します。

FILTERしたコンコーダンスでは、指定した部分が赤くなって表示されます。そのあとでFREQUENCY機能を使い、頻度を表示する位置を1. COLLOCATION (HIGHLIGHTED WORD)にします。

そうするとmadやcrazyが多いことが分かるはずです。

CONCORDANCE British National Corpus (BNC)

lemma drive • 15,596
138.84 per million tokens • 0.014%

Left context KWIC Right context

1 Written books a... the appointed Saturday in October, Micky met me at the railway station and drove me (in a fairly elderly Morris) right down the H

2 Written books a... certain . </s><s> Trains were not running, or if they were, they were driven by undergraduates or other surs . </s><s> In those days private ca

3 Written books a... western Japanese market . </s><s> * Southern Japanese fishermen drove about 1,000 dolphins back into the sea after having been informed by the gove

4 Written miscell... </s><s> In 1940 Scotland was 'invaded' by thousands of Polish soldiers, driven from their homeland by the German army . </s><s> The links they forget

FILTER

BASIC ADVANCED

Keep lines? containing

Query type simple lemma phrase word character CQL

CQL [tag="JJ"]

Default attribute? word

Range? Token Sentence Custom

-5 -4 -3 -2 -1 KWIC 1 2 3 4 5

FREQUENCY

BASIC ADVANCED ABOUT

Select an attribute and its position in the concordance: ?

word

left context right context

6 5 4 3 2 1 1. COLLOCATION (HIGHLIGHTED WORD) 1 2 3 4 5 6

Q69. BNCを利用し、Academicテキストで
mightを含む文をランダムに100文抽出しなさい。

Q69. BNCを利用し、Academicテキストでmightを含む文をランダムに100文抽出しなさい。

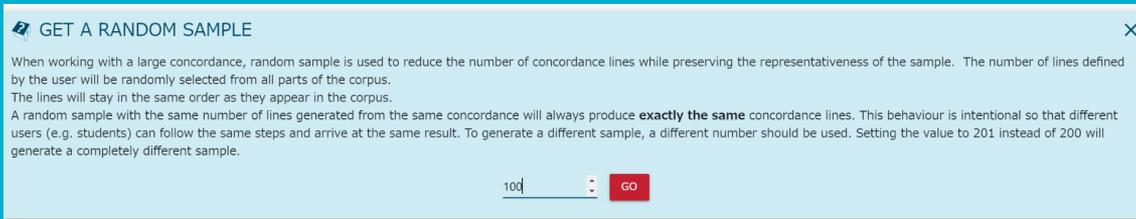
BNCでAcademicテキストに限定して検索するには、SubCorpusでWritten Academicを選択します。

ランダムに例を抽出する場合は、Get a random sampleを選択します。

すると右のような画面が出てくるので、抽出したいサンプル数を入力します。



The screenshot shows the BNC concordance interface. The search term is 'simple might', with 11,440 results. The interface includes a search bar, a filter for 'Written Academic', and a list of concordance lines. A red arrow points to the 'Get a random sample' button in the top right corner of the interface.



The screenshot shows the 'GET A RANDOM SAMPLE' dialog box. It contains the following text:

When working with a large concordance, random sample is used to reduce the number of concordance lines while preserving the representativeness of the sample. The number of lines defined by the user will be randomly selected from all parts of the corpus. The lines will stay in the same order as they appear in the corpus. A random sample with the same number of lines generated from the same concordance will always produce **exactly the same** concordance lines. This behaviour is intentional so that different users (e.g. students) can follow the same steps and arrive at the same result. To generate a different sample, a different number should be used. Setting the value to 201 instead of 200 will generate a completely different sample.

At the bottom, there is a text input field containing '100' and a red 'GO' button.

Q70. CQL: BNC 利用

"marriage"は"break up"する
か"break down"するか確認しなさい。

Q70. CQL: BNC 利用

"marriage"は"break up"するか"break down"するか確認しなさい。

CQL

```
[lemma="marriage"][]{0,2}[lemma="break"]  
[lemma="up|down"]
```

CQL **marriage + break** • 95
0.85 per million tokens • 0.000085%

Details **sentence**

1 He had been the injured party when his **marriage broke up** and, in spite of everything, it was all too clear that he still retained strong feelings for the woman he had married .
PP VHD VBN DT JJ NN WRB PPZ NN VVD RP CC , IN NN IN NN , PP VBD DT JJ INthat PP RB VVD JJ
NNS IN DT NN PP VHD VVN SENT

2 I am sure he has other purposes as well, it 's basically about a **marriage breaking down**
PP VBP JJ PP VHZ JJ NNS IN RB , PP VBZ RB IN DT NN VVG RP



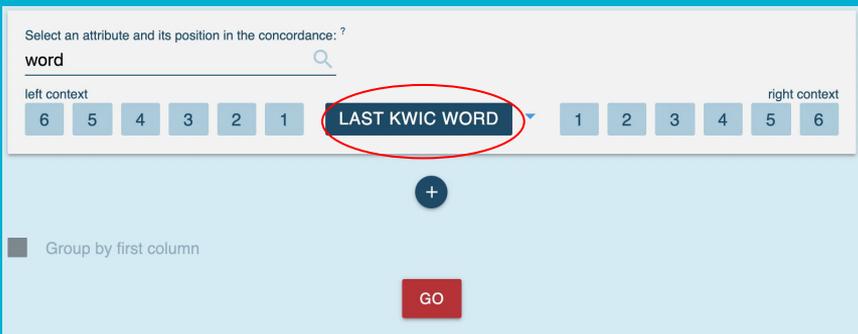
Select an attribute and its position in the concordance: ?
word

left context: 6 5 4 3 2 1 **LAST KWIC WORD** 1 2 3 4 5 6 right context

+

Group by first column

GO



(2 items, 95 total frequency)

	Word	Frequency ↓	Relative ?	% of conc. ?	
1	<input type="checkbox"/> up	54	0.48	56.84%	...
2	<input type="checkbox"/> down	41	0.36	43.16%	...



Q71. JEFLL_ORIGを用いてCEFRLレベルごとに1センテンス当たりの平均語数を出しなさい。

Q71. JEFLL_ORIGを用いてCEFRレベルごとに1文あたり平均語数を出しなさい。

CEFRレベルごとに文あたりの語数を出すためにはCEFRレベル別のサブコーパスの語数と、サブコーパスごとの文数を調べる必要があります。

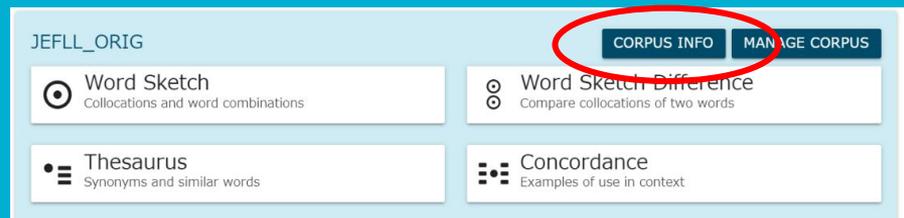
サブコーパスごとの語数はCorpus Infoで確認することができます。

サブコーパスごとの文数はコンコーダンスでサブコーパス指定をしたうえでCQL検索で<s>を検索します。

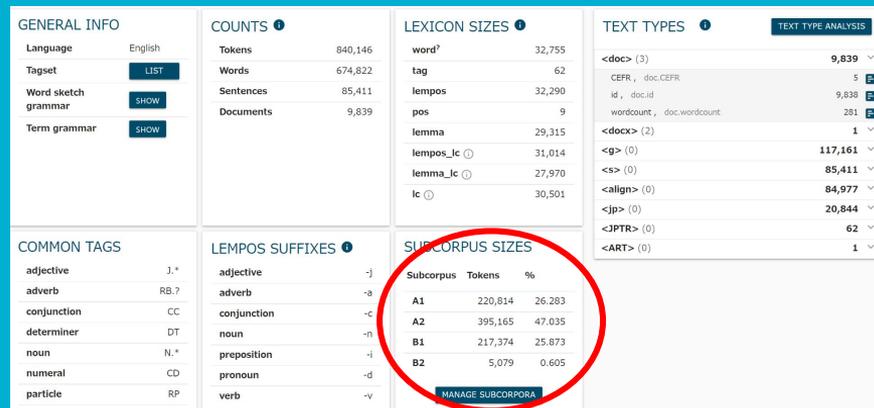
文数ごとの<s>というタグがついているのでこの数を調べることで文数を調べることができます。

A1の語数は220,814、文数は27,436なので、 $220,814 \div 27,436 \approx 8.05$ で、1文あたりの平均語数は約8.05となります。

同様に他のCEFRレベルも算出すると、CFERレベルが上がるごとに1文あたりの平均語数が増えることが分かります。



The screenshot shows the top navigation area of the JEFLL_ORIG interface. It includes a title 'JEFLL_ORIG' and two buttons: 'CORPUS INFO' and 'MANAGE CORPUS', both of which are circled in red. Below the buttons are four main menu items: 'Word Sketch' (Collocations and word combinations), 'Word Sketch Difference' (Compare collocations of two words), 'Thesaurus' (Synonyms and similar words), and 'Concordance' (Examples of use in context).



The screenshot displays the 'GENERAL INFO' section of the JEFLL_ORIG interface. It is divided into several panels: 'GENERAL INFO', 'COUNTS', 'LEXICON SIZES', 'TEXT TYPES', 'COMMON TAGS', 'LEMPOS SUFFIXES', and 'SUBCORPUS SIZES'. The 'SUBCORPUS SIZES' panel is circled in red and contains the following data:

Subcorpus	Tokens	%
A1	220,814	26.283
A2	395,165	47.035
B1	217,374	25.873
B2	5,079	0.605

Below the table is a button labeled 'MANAGE SUBCORPORA'.

Q72. BNC利用:
"recur"と"reoccur"のどちらがより
使われますか？

Q73. BNCでV + O + not to doのパターンをとる動詞を調べなさい。

Q73. BNCでV + O + not to doのパターンをとる動詞を調べなさい。

```
[tag="VV.?"][tag="DT"]?[tag="J.*"]?[tag="N.*|PP"
][lemma="not"][tag="TO"]
```

というCQLで検索します。

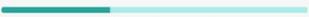
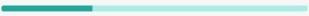
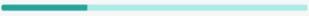
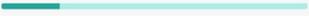
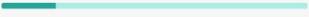
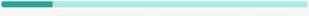
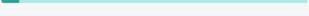
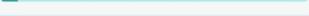
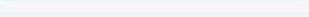
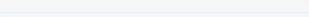
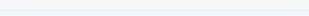
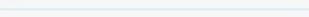
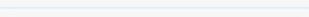
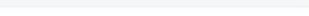
目的語の部分には普通の名詞がくる場合と代名詞がくる場合を想定したCQLを書きましょう。

また普通の名詞が目的語にくる場合は、それに限定詞が付いている可能性も考慮します。

tell, ask, warn, take, adviseなどが多い事が分かります。

ただしtakeの場合、take care not to というフレーズがほとんどで、調べたい構文パターンとは違うことがコンコードダンスを見ればわかります。

(175 items, 1,526 total frequency)

	Lemma	Frequency ↓	Relative ²	
1	<input type="checkbox"/> tell	369	3.28	 ...
2	<input type="checkbox"/> take	130	1.16	 ...
3	<input type="checkbox"/> warn	109	0.97	 ...
4	<input type="checkbox"/> ask	103	0.92	 ...
5	<input type="checkbox"/> advise	69	0.61	 ...
6	<input type="checkbox"/> urge	64	0.57	 ...
7	<input type="checkbox"/> persuade	61	0.54	 ...
8	<input type="checkbox"/> beg	48	0.43	 ...
9	<input type="checkbox"/> make	46	0.41	 ...
10	<input type="checkbox"/> trust	24	0.21	 ...
11	<input type="checkbox"/> force	20	0.18	 ...
12	<input type="checkbox"/> expect	19	0.17	 ...
13	<input type="checkbox"/> order	17	0.15	 ...
14	<input type="checkbox"/> encourage	16	0.14	 ...
15	<input type="checkbox"/> instruct	15	0.13	 ...
16	<input type="checkbox"/> want	15	0.13	 ...
17	<input type="checkbox"/> prefer	14	0.12	 ...
18	<input type="checkbox"/> will	13	0.12	 ...
19	<input type="checkbox"/> teach	13	0.12	 ...
20	<input type="checkbox"/> require	12	0.11	 ...

Q74. 後ろにby以外の前置詞が続く受動態表現を抽出しなさい。

Q74. 後ろにby以外の前置詞が続く受動態表現を抽出しなさい。

`[tag="VB.?"][tag="VVN"][tag="IN"&word!="by"]`

というCQLで検索します。

一般的にby以外の前置詞を使う受動態というbe surprised atなど感情を表すものに焦点を当てられがちですが、それ以外の表現でも多く使用されるものがあることが分かります。

(30,512 items, 400,442 total frequency)

		Lemma	Frequency ↓	Relative %	
1	<input type="checkbox"/>	be base on	4,569	40.67	...
2	<input type="checkbox"/>	be see as	3,993	35.54	...
3	<input type="checkbox"/>	be find in	3,886	34.59	...
4	<input type="checkbox"/>	be involve in	3,330	29.64	...
5	<input type="checkbox"/>	be regard as	2,966	26.40	...
6	<input type="checkbox"/>	be use in	2,809	25.00	...
7	<input type="checkbox"/>	be use for	2,658	23.66	...
8	<input type="checkbox"/>	be hold in	2,387	21.25	...
9	<input type="checkbox"/>	be give to	2,299	20.46	...
10	<input type="checkbox"/>	be use as	2,142	19.07	...
11	<input type="checkbox"/>	be make in	2,120	18.87	...
12	<input type="checkbox"/>	be associate with	2,038	18.14	...
13	<input type="checkbox"/>	be know as	1,943	17.29	...
14	<input type="checkbox"/>	be concern with	1,934	17.21	...
15	<input type="checkbox"/>	be send to	1,864	16.59	...
16	<input type="checkbox"/>	be describe as	1,779	15.84	...
17	<input type="checkbox"/>	be bear in	1,759	15.66	...
18	<input type="checkbox"/>	be show in	1,753	15.60	...
19	<input type="checkbox"/>	be see in	1,726	15.36	...
20	<input type="checkbox"/>	be include in	1,579	14.05	...

You are only allowed to access 1,000 items. [Get more](#)

Rows per page: 1-20 of 1,000

[|<](#) [<](#) [1](#) / 50 [>](#) [>|](#)

Q75. 文頭にOnly+副詞がくる文を検索し、どのような組み合わせが多いのか、またOnly+副詞の後ろは実際に倒置されているのか確認しなさい。

Q76. Parallel Concordance:

パラレルコンコーダンスでJEFLL_ORIGを用いて
become toを検索し、添削文ではどのように訂正さ
れているか確認しなさい。

Q76. パラレルコンコーダンスでEFLL_ORIGを用いてbecome toを検索し、添削文ではどのように訂正されているか確認しなさい。

パラレルコンコーダンスのCQLで

[lemma="become"][word="to"]を検索します。

学習者が書いたbecome toを使った文と、添削文を見比べてみると、単にbecomeでよいところをbecome to beとしているエラーなどが多いことが分かります。

Doc ID	Original Text	Corrected Text
doc#1192	<S> He said "Please become to before me" </s>	<S> He said , " Please come before me . " </s>
doc#2112	<S> He didn't become to young man . </s>	<S> He did n't become a young man . </s>
doc#3750	<S> When the butcher became like that, Wilbur became to 10 tears old and was in Ryugu-jou </s>	<S> When that happened to the butcher , Wilbur became 10 tears old and was in Ryugu-jou </s>
doc#3812	<S> And turu to kame became to medatai mono . </s>	<S> And the crane and the turtle became symbols of happiness . </s>
doc#3831	<S> He died, but he became to a turu . </s>	<S> He died but he became a crane . </s>
doc#4418	<S> He became to sell the box . </s>	<S> He had come to sell the box . </s>
doc#4473	<S> Because he want to become to be toreendiina shibui rojin . </s>	<S> because he wanted to become a trendy , cool man . </s>
doc#4744	<S> Because if I don't, I'll become to be hungry . </s>	<S> because , if I do n't , I 'll become hungry . </s>
doc#5060	<S> If there are a lot of money, I only become to be able to buy a lot of thing . </s>	<S> If I have a lot of money , I will only become able to buy a lot of things . </s>
doc#5375	<S> Nobody know the young man became to . </s>	<S> Nobody knew what the young man became . (</s><s> ?) </s>
doc#5477	<S> And , he became to a young man . </s>	<S> And he became a young man . </s>
doc#5589	<S> I didi n't want to do [JP:「 became to 」] , but I became to like last year . </s>	<S> I did n't want to do a bazaar but I came to like it , just as I did last year . (?) </s>
doc#5599	<S> A lot of people come and become to be glad . </s>	<S> A lot of people come and are gladdened . </s>
doc#5742	<S> We all thiked that No_one buy [JP:kakigori] but many people bought it and it became to be sold out ! ! </s>	<S> We all thought that no one would buy shaved ice , but many people bought it and it was sold out ! ! </s>

Q77. Parallel Concordance:

パラレルコンコーダンスでJEFLL_ORIGを用い、
<jp>タグを検索し、学習者が英語で表現できな
かったものが、添削文ではどのように表現されてい
るか確認しなさい、

Q77. パラレルコンコーダンスでJEFLL_ORIGを用い、<jp>タグを検索し、学習者が英語で表現できなかったものが、添削文ではどのように表現されているか確認しなさい、

パラレルコンコーダンスのCQLで

[<containing<jp/>

と入力します。

JEFLLでは学習者が英語で表現できないものは日本語で書くことが許されています。日本語で書かれた部分にはjp>のタグが付与されています。

上記のCQLで検索すればパラレルコンコーダンス上で日本語で表現されたものをKWIC表示ができます。FREQUENCY機能を使えば、日本語で表現されたものを頻度順に表示することもできます。

日本語で多く表現されたものが何かを確かめ、それが添削文でどのような英語で表現されているかを確かめれば、日本人学習者が英語にしづらいつと感じるものを、どのように英語にすればよいか学習できます。

PARALLEL CONCORDANCE

filter JEFLL_ORIG

CQL [containing=<jp/> 16,554 19,703.72 per million tokens 2% filter 貯金通帳 54 64.27 per million tokens 0.0064%

doc#567	<s> I will bring <jp> 貯金通帳 </jp> first. </s>	<s> I will take out my bankbook first </s>
doc#574	<s> I will bring <jp> 貯金通帳 </jp> and my friend's pictures. </s>	<s> I will take out my bankbook and my friend's pictures . </s>
doc#574	<s><jp> 貯金通帳 </jp> can buy something. </s>	<s> With my bankbook I can buy something . </s>
doc#575	<s> I will bring a new <jp> 貯金通帳 </jp> first. </s>	<s> I will take out a new bankbook first . </s>
doc#575	<s> My mother bought me that <jp> 貯金通帳 </jp> two weeks ago. </s>	<s> My mother bought me that bankbook two weeks ago . </s>
doc#575	<s> The <jp> 貯金通帳 </jp> is my <jp> 宝物 </jp>. </s>	<s> The bankbook is my treasure . </s>
doc#645	<s> I will bring <jp> 貯金通帳 </jp> first. </s>	<s> I will take out my bank book first . </s>
doc#653	<s> But <jp> やっぱり </jp> I will bring a <jp> 貯金通帳 </jp> first. </s>	<s> But actually I will take out my bankbook first . </s>
doc#676	<s> I will bring some <jp> 貯金通帳 </jp> first. </s>	<s> I will take out some bank book first </s>
doc#721	<s> I'm going to bring a <jp> サイフ、 </jp> a <jp> 貯金通帳 </jp> and some games. </s>	<s> I 'm going to take out a purse , a bankbook and some games . </s>
doc#1248	<s> Second I will bring <jp> 貯金通帳 </jp>. </s>	<s> Second , I will take out my bank book . </s>
doc#1270	<s> First, I will bring a <jp> 貯金通帳 </jp>. </s>	<s> First , I will take out my bank book . </s>
doc#1270	<s> I have <jp> 貯金通帳 </jp>. </s>	<s> I have a bank book . </s>
doc#1270	<s><jp> 貯金通帳 </jp> is <jp> たまっている10万円 </jp>. </s>	<s> In my bank account , I have saved 100000 yen . </s>
doc#1270	<s> I will bring my <jp> 貯金通帳 </jp> first. </s>	<s> I will take out my bank book first . </s>

Q78. Wordlist:

BNCでconから始まる名詞の頻度リストを出しなさい。

Q78. BNCでconから始まる名詞の頻度リストを出しなさい。

Wordlistでは、指定した文字列から始まる(または終わる、または含む)語を抽出して、頻度リストを作成することができます。

conから始まる名詞を指定するには、nounを選び、右側のstarting withを選びます。

すると文字を入力する欄が現われるので、そこにconと打ち込みます。

The screenshot shows the 'WORDLIST' interface for the 'British National Corpus (BNC)'. The 'ADVANCED' tab is active. In the 'find ?' section, 'words' is selected, and a dropdown menu is open with 'starting with' highlighted. The search input field contains 'con'. The 'noun' category is selected in the 'Text types ?' dropdown. The 'GO' button is visible at the bottom right.

Q79. Wordlist:

BNCで品詞を指定せずlemma単位で全ての語の頻度リストをWrittenとSpokenで出し、比較しなさい。

Q79. BNCで品詞を指定せずlemma単位で全ての語の頻度リストをWrittenとSpokenで出し、比較しなさい。

ADVANCEDタブでサブコーパスを指定をすることができます。

Spoken(話し言葉)に限定した、頻度リストを出す場合は、BNCではSpoken transcriptを、Written(書き言葉)に限定したい場合は(Written text)を選びます。

WORDLIST British National Corpus (BNC)

BASIC ADVANCED ABOUT

find ? words all
lemmas starting with
adjective ending with
adverb containing
conjunction matching regex
noun from this list:
preposition
pronoun

Exclude these words:

Include nonwords ?

A = a ?

Frequency min ? 5 Frequency max ? 0

result format
 Simple list ?
 Display as ?

Subcorpus
Spoken transcripts

Subcorpus and Text types cannot be combined
Text types ? ▾

GO

Q80. Wordlist:

BNCでconかcomから始まる形容詞の頻度リストを出しなさい。

Q80. BNCでconかcomから始まる形容詞の頻度リストを出しなさい。

matching regexを選ぶと、正規表現に一致した語を抽出した頻度リスト出すことができます。

co[nm].* と正規表現を使って書くと、conかcomから始まる語を全て指定できます。

[]の記号は中の文字のどれか1つを指定します。[nm]だとnかmのどちらかを表します。

.(ドット)の記号は任意の一語を意味します。*(アスタリスク)は直前の文字を0回以上繰り返すという意味です。

The screenshot shows the BNC search interface with the following configuration:

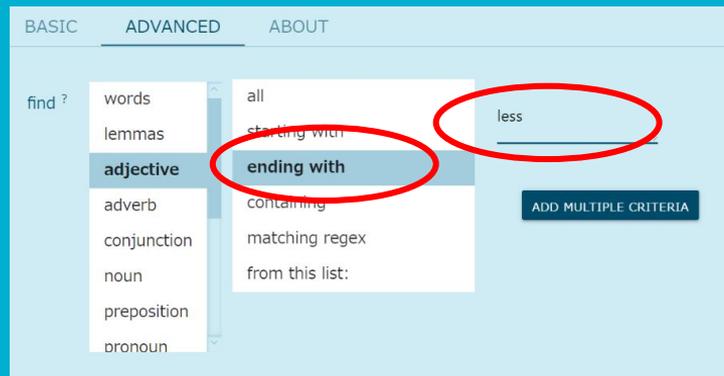
- Tab: **ADVANCED**
- find ? dropdown: **matching regex** (circled in red)
- Search query: **co[nm].*** (circled in red)
- Category dropdown: **adjective**
- Frequency min: **5**
- Frequency max: **0**
- result format: **Simple list ?** (selected)
- Subcorpus: **none (the whole corpus)**
- GO button: **GO**

Q81. Wordlist:
BNCでlessで終わる形容詞のリス
トを出しなさい。

Q81. BNCでlessで終わる形容詞のリストを出しなさい。

品詞をadjectivesに指定し、ending withを選び、lessを入力します。

すると右のようにlessで終わる形容詞のリストを簡単に出すことができます。



WORDLIST

adjective (248 items | 42,763 total frequency)

	Lemma	Frequency ? ↓	Frequency Per Million ? ↓	DOCF ?
1	less	18,998	169.10	2,989 ...
2	endless	1,512	13.46	825 ...
3	useless	1,264	11.25	755 ...
4	helpless	790	7.03	442 ...
5	hopeless	707	6.29	461 ...
6	homeless	697	6.20	288 ...
7	meaningless	641	5.71	443 ...
8	harmless	630	5.61	426 ...
9	ruthless	627	5.58	383 ...
10	countless	617	5.49	435 ...

Q82. Wordlist:

BNCでnaturalとその派生語
(語の中にnaturalを含む単語)
の頻度リストを出しなさい。

Q82. BNCでnaturalとその派生語(語の中にnaturalを含む単語)の頻度リストを出さない。

語の中に特定の文字列を含む語を抽出して、頻度をリストを出すためには、containingを選び、文字列を指定します。

The screenshot shows the BNC search interface with the following settings:

- Navigation tabs: BASIC, **ADVANCED**, ABOUT
- find ? dropdown: words, **lemmas**, adjective, adverb, conjunction, noun, preposition, pronoun
- Search criteria dropdown: all, starting with, ending with, **containing**, matching regex, from this list:
- Search term: natural
- Button: ADD MULTIPLE CRITERIA
- Exclude these words:
- Include nonwords:
- A = a:
- Frequency min: 5, Frequency max: 0
- result format: Simple list, Display as
- Subcorpus: none (the whole corpus)
- Text types: Text types ?
- GO button

Q83. Wordlist:

BNCで、サブコーパスをWritten academicに指定し、最大頻度200、最低頻度50に指定して、形容詞のリストを出しなさい。

Q83. BNCで、サブコーパスをWritten academicに指定し、最大頻度200、最低頻度30に指定して、形容詞のリストを出しなさい。

抽出する語の最大頻度と最低頻度を指定してワードリストを抽出することができます。

Frequency minを50にし、Frequency maxを200にします。

SubcorpusでWritten Academicにすれば、アカデミックテキストに絞って抽出することができます。

The screenshot shows the BNC search interface with the following settings:

- find ? : adjective
- Frequency min ? : 50
- Frequency max ? : 200
- subcorpus ? : Written Academic

Additional options visible include:

- Exclude these words:
- Include nonwords ?
- A = a ?
- result format: Simple list ? Display as ?

Buttons for "GO" and "lock" are also visible.

Q84. Starting with "What"にしてspoken striptと academic writingで質問の使い方が大きく異なることを確認する。

Q84. Starting with "What"にしてspoken striptとacademic writingで質問の使い方が大きく異なることを確認する。

サブコーパス間で比べれば、ある語の振る舞い方のレジスター別の違いを具体的なフレーズから見てとることができます。

CHANGE CRITERIA

BASIC ADVANCED ABOUT

N-gram length? 2 3 4 5 6

Attribute? word

Frequency min? 5 Frequency max? 0

Subcorpus? Spoken transcripts

Additional criteria

- all
- starting with letters
- ending with letters
- containing letters
- starting with word**
- containing word
- ending with word
- matching a regular expression

what

Subcorpus and Text types cannot be combined

Text types?

Q85. N-gram: BNC利用

ICNALEを使い、CEFRLレベルA2とB2の学習者と母語話者で、ofを含む4-gramをそれぞれ調べ、比較しなさい。

Q85. ICNALEを使い、CEFRレベルA2とB2の学習者と母語話者で、ofを含む4-gramをそれぞれ調べ、比較しなさい。

n-gram検索でも、Subcorpusを指定して検索をすることができます。

ICNALEはSketch Engine上で東京外国語大学内アカウントであれば、デフォルトでサブコーパスが利用できます。

CEFRレベル別に検索するにはサブコーパス指定をして検索をします。母語話者は「Country_NativeSpeaker」を選びます。

検索画面の左上のところでも、サブコーパスを簡単に切り替えられます。

BASIC ADVANCED ABOUT

N-gram length ? 2 3 4 5 6

Attribute ? word

Frequency min ? 3 Frequency max ? 0

Subcorpus ? CEFR_A2

Additional criteria ? containing word

of

N-GRAMS ICNALE_WE

country_NativeSpeaker x 4-grams, word (items: 124, total frequency: 571)

Word	Frequency ?	Word	Frequency ?
1 the value of money	17 ...	18 some of my friends	6 ...
2 is one of the	17 ...	19 and a lot of	6 ...
3 as a result of	13 ...	20 of my friends have	6 ...
4 a little bit of	12 ...	21 smell of cigarette smoke	6 ...
5 a lot of people	11 ...	22 have a lot of	6 ...

Q86. 特徴語抽出: ICNALE_WEを選択後、Subcorpusを利用して日本人学習者と英語母語話者間の特徴語を抽出しなさい。

Q86.

KEYWORDS

ICNALE_WE

Get more space

Recompile your corpus.
This corpus uses outdated term extraction configuration (term grammar). Recompile your corpus with the latest term grammar for better results.

BASIC **ADVANCED** ABOUT

Focus subcorpus ?
country_Japan

Reference corpus ?
ICNALE_WE

Reference subcorpus ?
country_NativeSpeaker

Focus on ?
rare 1 common

Minimum frequency ? 1
Maximum frequency ? 0

Maximum items ? 1000

- A = a ?
- At least one alphanumeric ?
- Only alphanumeric ?
- Include nonwords ?
- Exclude these words ?
- From list

Identify keywords
 Identify terms
 Identify n-grams

Keywords settings
Attribute ?
lemma

Terms settings
Matching regex ?
*

N-grams settings
Attribute ?
word

Word	Word
1 smoker	11 part-time-job
2 smoking	12 ban
3 non-smoker	13 no-smoke
4 nonsmoker	14 tobacco
5 part-time	15 smell
6 non-smoke	16 nicotine
7 nonsmoking	17 passive
8 smoke	18 dislike
9 cigarette	19 demerit
10 restaurant	20 no-smoker

母語話者が余り使わない
“demerit”が特徴的として
抽出されている。

Q87. 特徴語抽出: ICNALE_WEを用いて、topic_part-time jobと比較して、topic_smokingの特徴的な語彙を抽出しなさい。

Q87. ICNALE_WEを用いて、topic_part-time jobと比較して、topic_smokingの特徴的な語彙を抽出しないでください。

学習者コーパスでは、topicの制約により頻度が多い語が出てきてしまいます。topicによってよく使われる語を抽出すれば、例えばそれを取り除いてtopicによる語彙的な影響を抑えた検索などができるようになります。

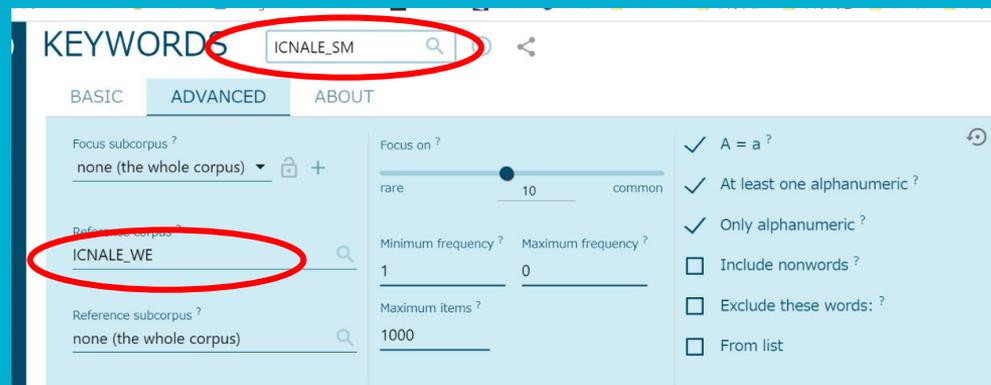
KeyWords検索を使えばtopicに特徴的な語彙を抽出することができます。

The screenshot shows the 'CHANGE CRITERIA' interface with three tabs: BASIC, ADVANCED, and ABOUT. The 'ADVANCED' tab is selected. The 'Focus subcorpus' dropdown menu is set to 'topic_Smoking' and is circled in red. Below it, the 'Reference corpus' is set to 'ICNALE_WE'. The 'Reference subcorpus' dropdown menu is set to 'topic_Part-time_job' and is also circled in red. To the right, there are several search criteria: 'Focus on' with a slider set to 100 (between 'rare' and 'common'), 'Minimum frequency' set to 1, 'Maximum frequency' set to 0, and 'Maximum items' set to 1000. On the far right, there are checkboxes for 'A = a', 'At least one alphanumeric', 'Only alphanumeric', 'Include nonwords', 'Exclude these words', and 'From list'. The first three are checked, while the last three are unchecked.

Q88. 特徴語抽出: ICNALE_SMを用いて、ICNALE_WEと比較したときの特徴的な語彙を抽出しなさい。

Q88. ICNALE_SMを用いて、ICNALE_WEと比較したときの特徴的な語彙を抽出しなさい。

ICNALEのSMとWEを比較すれば、学習者の書き言葉と話し言葉で特徴的な語彙を抽出することができます。



KEYWORDS ICNALE_SM

BASIC ADVANCED ABOUT

Focus subcorpus? none (the whole corpus)

Reference corpus? ICNALE_WE

Reference subcorpus? none (the whole corpus)

Focus on? rare 10 common

Minimum frequency? 1 Maximum frequency? 0

Maximum items? 1000

- A = a?
- At least one alphanumeric?
- Only alphanumeric?
- Include nonwords?
- Exclude these words: ?
- From list

	Word	Frequency per million?			Word	Frequency per million?		
		Focus	Reference			Focus	Reference	
1	uh	7,916.35	0.00 ...	18	actually	1,105.60	276.97 ...	
2	um	3,068.14	0.00 ...	19	hello	33.99	1.40 ...	
3	yeah	688.77	4.20 ...	20	disagree	951.75	251.09 ...	
4	uhh	293.40	0.00 ...	21	um-hmm	25.05	0.00 ...	
5	uhm	257.62	0.00 ...	22	hmm	26.84	0.70 ...	
6	umm	230.78	0.00 ...	23	smo	28.62	1.40 ...	
7	ah	198.58	0.00 ...	24	eh	23.26	0.00 ...	
8	okay	434.73	29.38 ...	25	guess	127.02	31.47 ...	
9	etcetera	76.93	0.70 ...	26	ex	32.20	2.80 ...	

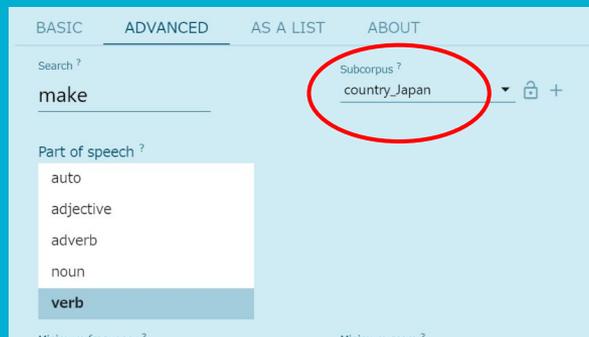
上級編

Q89. ICNALE_WEでmakeの目的語になる名詞を、subcorpus指定でJapanとPakistanでそれぞれ調べ、比較しなさい。

Q89. ICNALE_WEでmakeの目的語になる名詞を、subcorpus指定でJapanとPakistanでそれぞれ調べ、比較しなさい。

ICNALE_WE(Written Essays)では、学習者の国籍別にサブコーパス指定をできるようにしてあります。

サブコーパス設定をして、日本人英語学習者とパキスタン人英語学習者におけるmakeの使用みると、明らかな違いが見てとれます。

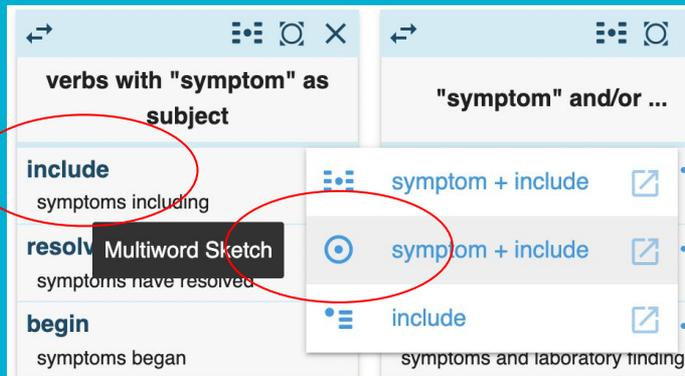
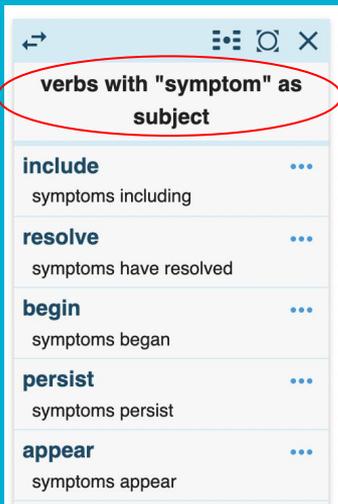
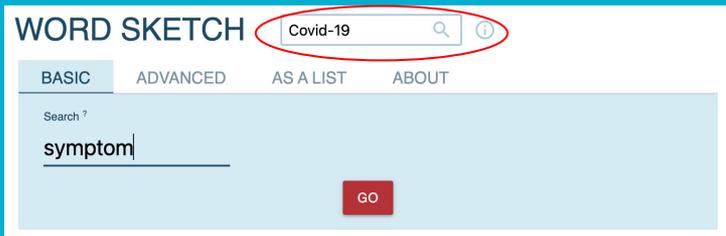


The screenshot shows the WORD SKETCH interface for the search term 'make' as a verb. The subcorpus is 'country_Japan'. The analysis is divided into four columns: modifiers of "make", objects of "make", subjects of "make", and prepositional phrases. The 'objects of "make"' column is circled in red.

modifiers of "make"	objects of "make"	subjects of "make"	prepositional phrases
easily 2 7.7 ...	friend 79 11.4 ...	smoke 34 9.6 ...	"make" by ... 0.4% ...
only 8 7.0 ...	effort 14 9.5 ...	tobacco 7 9.2 ...	"make" of ... 0.3% ...
often 2 6.9 ...	relation 10 9.2 ...	cook 2 8.5 ...	"make" because ... 0.3% ...
much 2 6.3 ...	mistake 10 9.1 ...	cigarette 12 8.4 ...	
also 4 5.0 ...	rule 12 8.9 ...	smell 3 8.4 ...	
not 22 4.9 ...	money 68 8.8 ...	air 2 8.0 ...	
	wall 7 8.7 ...	Japan 2 7.8 ...	
	people 39 8.6 ...	school 2 7.8 ...	
	taste 7 8.4 ...	experience 3 7.6 ...	
	relationship 6 8.2 ...	smoking 34 7.6 ...	
	life 13 8.2 ...	government 3 7.1 ...	
	space 6 8.2 ...	reason 3 7.1 ...	
	use 5 8.0 ...	place 2 6.9 ...	
	dish 5 7.9 ...	restaurant 8 6.7 ...	
	smoker 15 7.9 ...	job 12 6.4 ...	
	seat 5 7.8 ...	person 2 6.1 ...	
	nonsmoker 4 7.7 ...	smoker 3 5.5 ...	
	matter 4 7.7 ...	people 5 4.7 ...	
	arc 8 7.7 ...	student 2 3.6 ...	

Q90. Word Sketch:
"symptom"と"include"の組み合わせを使ってCovid19の代表的な症状を確認しなさい。

Q90. "symptom"と"include"の組み合わせを使ってCovid19の代表的な症状を確認しなさい



Q91. Word Sketch:

句動詞"carry out"と同じ意味を持つ動詞はどのようなものがありますか？"carry out"の目的語を用いて確認しなさい。

Q91. 句動詞"carry out"と同じ意味を持つ動詞はどのようなものがありますか？"carry out"の目的語を用いて確認しなさい

WORD SKETCH British National Corpus (BNC)

BASIC ADVANCED AS A LIST ABOUT

Search? carry out Subcorpus? none (the whole corp...)

Part of speech? auto

verb

Minimum frequency? auto Minimum score? 0

GO

objects of "carry out"

survey	49	8.9 ...
research	59	8.7 ...
repair	24	8.5 ...
experiment	31	8.4 ...
study	65	8.4 ...
test	60	8.3 ...
work	175	8.1 ...
investigation	25	7.9 ...
analysis	27	7.7 ...
operation	32	7.5 ...
assessment	17	7.4 ...
check	13	7.4 ...

最も共起スコアの高い"survey"を使う

WORD SKETCH British National Corpus (BNC)

BASIC ADVANCED AS A LIST ABOUT

Search? survey Subcorpus? none (the whole corp...)

Part of speech? auto

noun

Minimum frequency? auto Minimum score? 0

GO

"survey"をもう一度
WordSketchで検索

verbs with "survey" as object

conduct	245	10.6 ...
survey conducted		
commission	43	9.0 ...
undertake	73	8.9 ...
publish	62	7.8 ...
survey published		
complete	43	7.4 ...
carry	50	6.9 ...
survey was carried out		
mount	11	6.5 ...
quote	11	6.4 ...
design	20	6.3 ...
survey designed to		
sponsor	7	6.2 ...
conclude	7	6.2 ...
organise	13	6.2 ...

Q92. CQL検索で動詞のserve + 名詞のパターンを、名詞の前に限定詞や形容詞や所有格が付く可能性も考慮して抽出しなさい。

Q92. CQL検索で動詞のserve + 名詞のパターンを、名詞の前に限定詞や形容詞や所有格が付く可能性も考慮して抽出しなさい。

[lemma="serve"&tag="V.*"][tag="N.*"]というCQLでは、一応動詞のserve+その目的語を抽出できたこととなりますがserveの直後に名詞が来ているパターンなので、目的語となる名詞に限定詞(he, a, some など)や形容詞や所有格がついている可能性を無視しています。ここではserveの目的語となる名詞を正確に抽出できたとは言えません。限定詞や形容詞が場合も考慮するためには、

[lemma="serve"&tag="V.*"][tag="DT|J.*|[CNP].*Z"]?[tag="N.*"]

という検索式を書きます。(DTは限定詞、J.*は比較級・最上級を含む形容詞、[CNP].*Zは所有格を指定しています)

?をつけると、その前の語があってもなくてもよいという意味になります。従って上のCQLの[tag="DT|J.*|[CNP].*Z"]?の部分は、限定詞が形容詞か所有格があってもなくてもよいということになります。

CONCORDANCE British National Corpus (BNC)

BASIC ADVANCED ABOUT

Query type ?

- simple
- lemma
- phrase
- word
- character
- CQL**

CQL

[lemma="serve"&tag="V.*"][tag="DT|J.*|[CNP].*Z"]?[tag="N.*"]

Insert [] { } < > ~ & ¥ | ~ # TAGS

CQL BUILDER

Default attribute?

lemma

(1,121 items, 3,725 total frequency)

	Lemma	Frequency ↓	Relative ?	
1	<input type="checkbox"/> purpose	237	2.11	...
2	<input type="checkbox"/> interest	127	1.13	...
3	<input type="checkbox"/> notice	90	0.80	...
4	<input type="checkbox"/> sentence	90	0.80	...
5	<input type="checkbox"/> life	87	0.77	...
6	<input type="checkbox"/> need	85	0.76	...
7	<input type="checkbox"/> food	55	0.49	...
8	<input type="checkbox"/> customer	55	0.49	...
9	<input type="checkbox"/> time	54	0.48	...
10	<input type="checkbox"/> dish	53	0.47	...
11	<input type="checkbox"/> apprenticeship	52	0.46	...
12	<input type="checkbox"/> people	47	0.42	...
13	<input type="checkbox"/> community	41	0.36	...
14	<input type="checkbox"/> drink	40	0.36	...
15	<input type="checkbox"/> function	37	0.33	...
16	<input type="checkbox"/> meal	36	0.32	...
17	<input type="checkbox"/> member	34	0.30	...
18	<input type="checkbox"/> country	34	0.30	...
19	<input type="checkbox"/> prison	33	0.29	...

Q93. CQL検索でonから始まりbasisで終わる
4~6単語連鎖を、文をまたぐものを除いて検索
しなさい。

Q93. CQL検索でonから始まりbasisで終わる4~6単語連鎖を、文をまたぐものを除いて検索しなさい。

[lemma="on"][[{2,4}[lemma="basis"]within</s>

というCQLを書きます。

within</s>をつけると、sentence内であることを指定でき、文をまたいでしまうものを取り除くことができます。

The screenshot shows the CONCORDANCE search interface for the British National Corpus (BNC). The search query is `[lemma="on"][[{2,4}[lemma="basis"]within</s>`, which has returned 4,434 results. The results are displayed in a table with four columns: Details, Left context, KWIC, and Right context. The KWIC column highlights the search results, showing phrases like "on a semi-itinerant basis", "on a daily basis", "on a regular basis", "on a pilot basis", "on a continuous or modular basis", "on a regular basis", "on a part-time basis", "on a very regular basis", "on a commercial basis", "on a part-time basis", "on a fees-only basis", "On a comparative basis", "on a convertible rouble basis", "on a sustainable yield basis", "on a sustainable basis", "on an ad hoc basis", "on a comparative basis", and "on a bid to bid basis".

Q94. CQL検索で、動詞のcontain+ 名詞のパターンを、名詞の前に複数個、限定詞や形容詞や所有格などが付く可能性も考慮して抽出しなさい。

Q94. CQL検索で、動詞のcontain+ 名詞のパターンを、名詞の前に複数個、限定詞や形容詞や所有格などが付く可能性も考慮して抽出しなさい。

```
[lemma="contain"&tag="V.*"] [tag="DT|J.*|[CNP].*Z"] {0,2} [tag="N.*"]
```

というCQLを書きます。

[tag="DT|J.*|[CNP].*Z"]{0,2}の部分は、「限定詞か形容詞か所有格が0~2個」という意味になります。containの目的語となる名詞の前に2つくらいまでであれば、限定詞や形容詞がつくことは大いにあり得るので、この方がcontain+その目的語のパターンの抽出としては妥当性がありそうです。

検索結果のConcordanceラインを見ればわかりますが、しっかりとcontain+その目的語のパターンを、目的語の前に限定詞や形容詞が0個から2個ついているものを含めて抽出できています。

CONCORDANCE British National Corpus (BNC)

BASIC ADVANCED ABOUT

Query type ⓘ

- simple
- lemma
- phrase
- word
- character
- CQL**

CQL

```
[lemma="contain"&tag="V.*"] [tag="DT|J.*|[CNP].*Z"] {0,2} [tag="N.*"]
```

Insert [] { } < > ~ & ¥ | ~ # TAGS

CQL BUILDER

Default attribute ?

lemma

Q95. V +所有格 + way +前置詞or副詞のパターン(いわゆるway構文)を検索し、頻度の高い実現形を調べなさい。その際way構文ではないものが極力含まれないように工夫して検索しなさい。

Q95. V + 所有格 + way + 前置詞or副詞のパターン(いわゆるway構文)を検索し、頻度の高い実現形を調べなさい。その際way構文ではないものが極力含まれないように工夫して検索しなさい。

まずは、`[tag="V:*"][tag=":*Z"][word="way"][tag="IN|RB"]`というCQLを書いてみます。これで大体は、抽出したいものが抽出できていますが、動詞の部分にbe動詞がくるものもそれなりの数があります。しかし実際の文を見てみると分かるようにbe動詞を使ったものはway構文ではありません。なのでbe動詞を除くという工夫が必要です。

そのために、`[tag="V:*"&tag!="VB.*"][tag="[CNP].*Z"][word="way"][tag="IN|RB"]`というCQLを書きます。

[属性="値"]というCQLの基本に!`!`をつけ足して、[属性!="値"]または[!属性="値"]とすると、指定した値を除くという意味になります。従って上記のCQLの `tag!="VB.*"` の部分はbe動詞を除くという意味です (VB.*は活用形も含むbe動詞を指定するタグ)。

way構文は移動を表す構文であり、wayの後ろには場所や方向を表す前置詞や副詞が通常は来ます。検索結果をみるとofが来ているものがありますが、これはway構文ではありません。なのでofがくるものを取り除き、

`[tag="V:*"&!tag="VB.*"][tag="[CNP].*Z"][word="way"][tag="IN|RB"&word!="of"]`というCQLを書くことでさらに正確に抽出することができます。

このように検索結果を見て、不必要であると思われるものを取り除いて検索することで、より正確な検索ができるようになります。

Q96. someone to talk toのように、不定詞の中身が前置詞で終わるパターンを抽出しなさい。

Q96. someone to talk toのように、不定詞の中身が前置詞で終わるパターンを抽出しなさい。

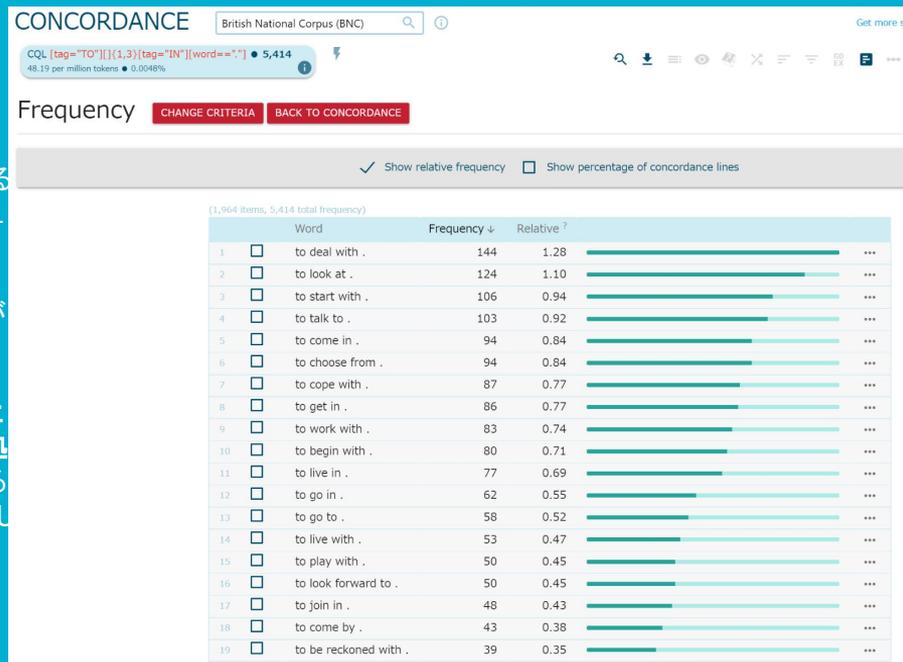
`[tag="TO"][[1,3][tag="IN"]][word==""]`

というCQLを書きます。不定詞の中身が前置詞で終わるパターンを厳密に抽出するのは難しいですが、不定詞のtoの後に、数単語入り、文末に前置詞がくるといふQLを書くことで便宜的に抽出することができます。

前置詞の後にピリオドがくるということを指示するQLを書くことで、文末に前置詞がくるパターンを指定できます。

上記の`[word==""]`はピリオドを表しています。今までと違い`=`が2つ続いていることに注目してください。こうすることで”の中身が正規表現ではなく単なる文字列として処理されます(正規表現のエスケープ。「.」の記号は通常では任意の1文字を指定する正規表現として処理されるので、正規表現としてではなく単なるピリオドとして処理したい場合はこのように`=`の記号を2つ重ねます。

通常の正規表現のエスケープ記号と同じようにバックスラッシュを使って、`[word="\."]`ともできます。



Q97. take ~ for grantedの表現で、～の部分にくる名詞もしくは代名詞を調べなさい。

Q97. take ~ for grantedの表現で、～の部分にくる名詞もしくは代名詞を調べなさい。

`[tag="N.*|PP"]within[lemma="take"]][1,3][word="for"]`
`[word="granted"]`

というCQLを書きます。within</>は文内であることを指定するCQLでしたが、withinの後ろには自由にCQLを書くことができます。

上記のCQLではtakeとfor grantedの間に何か1語～3語きて、その部分に入ってくる名詞もしくは代名詞を指定しています。要するにtakeとfor grantedの間に入ってくる名詞か代名詞を抽出していることになります。

CONCORDANCE British National Corpus (BNC)

CQL `[tag="N.*|PP"]within[lemma="take"]][1,3][word="for"] [word="granted"]`
3.79 per million tokens • 0.00038%

	Details	Left context	KWIC	Right context
51	<input type="checkbox"/>	Written books a... Meg put away in the kitchen without examining . . .	</s><s> She took them	for granted , of course . . . Arthur brought them every week
52	<input type="checkbox"/>	Written books a... high as second behind Laura Davies . . .	</s><s> Just as everyone took it	for granted that the young Scot was on the point of fulfilling the potentia
53	<input type="checkbox"/>	Written books a... the current league champions and Jim Nelson's men ca n't afford to take	anything	for granted . . . Only last Sunday the Royal Liver League cham
54	<input type="checkbox"/>	Written books a... hange in media will continue at this pace for some years and we can take	nothing	for granted . . . Media owners will need to constantly refocus
55	<input type="checkbox"/>	Written books a... eated with the utmost respect . . . We have learned to take	nothing	for granted . . . The Schweppes Cup is a long hard haul ,
56	<input type="checkbox"/>	Spoken context... ce in the terms of our employment . . . No longer do we take	it	for granted as meaning that we get a job in a company , or an indus
57	<input type="checkbox"/>	Written miscell... nd professionalism , but this trial has shown that we can afford to take	nothing	for granted . . . Mr. Ashby </s><s> I think that the whole I
58	<input type="checkbox"/>	Written miscell... > The crucial lesson for all political parties must be that we can not take	growth	for granted , and above all , we should not plan in advance how to spen
59	<input type="checkbox"/>	Written miscell... convictions for serious crimes . . . The police service takes this	service	for granted , and that is how it should be . . . My right hon.
60	<input type="checkbox"/>	Written miscell... ving stresses of traffic on them . . . They tend to take existing	tranquility	for granted - until it has gone . . . ' Goodbye , Britain ?
61	<input type="checkbox"/>	Written books a...) explain the existence and internal structure of firms rather than to take	them	for granted . . . In doing so , it rightly emphasizes the need f
62	<input type="checkbox"/>	Written books a... tner on the tiller . . . He seemed to be taking her continued	help	for granted now - issuing his orders . . . She ought to feel rei
63	<input type="checkbox"/>	Written books a... es of history formerly the preserve of historians , and can no longer take	history	for granted . . . History is not some unmediated story of eve
64	<input type="checkbox"/>	Written books a... ch and the themes to which they devote most of their attention take this	distinction	for granted . . . In so far as this is still true it is a hangov
65	<input type="checkbox"/>	Written books a... im . . . The United Nations Declaration of Human Rights takes	it	for granted not merely that all individual men are members of a single
66	<input type="checkbox"/>	Written books a... </s><s> Cutting a whole sequence of further corners I shall now take	it	for granted that the total process by which we habitually segment and c
67	<input type="checkbox"/>	Written books a... ratic agents of the established political order and that they always took	it	for granted that hierarchy is part of the natural order of things </s
68	<input type="checkbox"/>	Written books a... scribe their observations in language which takes the anthropomorphic	analogy	for granted . . . They regularly assert that the significance of

Q98. it is 形容詞 thatという形式主語構文で使われる形容詞の頻度リストを出しなさい。その際にbe動詞の時制や、助動詞が付く可能性も考慮しなさい。

Q98. it is 形容詞 that という形式主語構文で使われる形容詞の頻度リストを出しなさい。その際I be 動詞の時制や、助動詞が付く可能性も考慮しなさい。

`[tag="J.*"]within[lemma="it"][tag="MD"]?[tag="VH.*"]?[tag="RB.?"]?[tag="VB.?"]?[tag="J.*"]?[tag="IN/that"]`

というCQLで検索します。

withinの後ろの部分は、be動詞に助動詞がついたり現在完了形になる可能性を考慮しています。そのためtag="MD"]?(助動詞があってもなくもよいという意味や[tag="VH.*"]?(haveとその変化形があってもなくもよいという意味をつけています。また形容詞の前に副詞(notを含む)が付く場合も考慮して[tag="RB.?"]?というCQLをつけています。withinの後ろのこの構造で大体it 形容詞that というパターンを時制等も考慮して抽出できます。

withinの前に[tag="J.*"]を付けることでwithin内の[tag="J.*"]の部分にくるものを抽出することができます。

Q99. ランダムサンプルとアノテーションの機能を用いて、recentlyがどの時制でどの割合で使われているか調べなさい。

Q99. ランダムサンプルとアノテーションの機能を用いて、recentlyがどの時制でどの割合で使われているか調べなさい。

recentlyという語がどの時制でどの割合で使われているかは、すぐには調べられないのでランダムサンプルとアノテーション機能を使います。

まずrecentlyはコンコーダンスで検索しランダムサンプルの機能で例えば100文、recentlyを含む文を無作為に抽出します。

次にアノテーションモードを選びます。recentlyは現在完了か過去形のどちらかで使うと思われるので、ひとまずMANAGE ANNOTATIONSのところから、past tenseとpresent perfectの2つをLabelとして用意します。

次のコンコーダンスラインを一つ一つ見てrecentlyがpast tenseかpresent perfectかどちらで使われているか判断しLabelを選択します。

こうすることでrecentlyを含む文全てを確かめることはできませんが、無作為に抽出した100例からrecentlyがどれくらいの割合でどの時制で用いられているかを推定することができます。

CONCORDANCE British National Corpus (BNC)

simple recently • 12,178 (108.4 per million tokens • 0.011%) Sample 100 • 100 (0.89 per million tokens • 0.00089%)

Left context KWIC Right context

1 Written books a... y the Europeans and South Americans . </s><s> Indeed , recently th... ch Government suggested that in addition to its measure to re...

CONCORDANCE ANNOTATION MODE

The concordance annotation lets you categorize or add labels to concordance lines. It does not annotate the source text of the corpus. Your annotations cannot be downloaded but can be used in analysing the concordance.

Annotation mode is now active. Annotation of each concordance line is saved immediately. Closing this mode will preserve all annotations which can be accessed again from the dashboard.

Label	Filter	Frequency
past tense	<input checked="" type="radio"/> <input checked="" type="radio"/>	0
present perfect	<input checked="" type="radio"/> <input checked="" type="radio"/>	0

MANAGE ANNOTATIONS

SORT BY LABELS

EXIT ANNOTATION MODE

CONCORDANCE ANNOTATION MODE

The concordance annotation lets you categorize or add labels to concordance lines. It does not annotate the source text of the corpus. Your annotations cannot be downloaded but can be used in analysing the concordance.

Annotation mode is now active. Annotation of each concordance line is saved immediately. Closing this mode will preserve all annotations which can be accessed again from the dashboard.

Label	Filter	Frequency
past tense	<input checked="" type="radio"/> <input checked="" type="radio"/>	39
present perfect	<input checked="" type="radio"/> <input checked="" type="radio"/>	35

MANAGE ANNOTATIONS

SORT BY LABELS

EXIT ANNOTATION MODE

Details Left context KWIC Label Right context

51 Written books a... y the Europeans and South Americans . </s><s> In China , recently past tense foreign journalists were forced by fevered local fans to agree...

52 Written books a... essential aspect of the discourse situation . </s><s> More recently present perfect till , Halliday's functional model of language acknowledges th...

53 Written books a... rian , Marshal Petain . </p></s><s><p> Radio had only recently present perfect developed into a mass medium for news and entertainment...

54 Written books a... ssor of Classical Art and Archaeology at Oxford University , recently past tense locked an audience of classicists by claiming that : ' The...

Q100. Parallel Concordance:

パラレルコンコーダンスで、ICNALE_EEを用い、学習者が現在完了形を使うべきでないところで使っているエラーを含む文を抽出しなさい。

Q100. パラレルコンコーダンスで、ICNALE_EEを用い、学習者が現在完了形を使うべきでないところで使っているエラーを含む文を抽出しなさい。

パラレルコンコーダンスのADVANCED検索では、通常のコンコーダンスと同じように、様々なQuery typeで検索することができ、CQLも使うことができます。

さらに対応する文に特定のQueryを含む(or含まない)を指定して、結果を絞ることができます。

まず、CQLで[word="have|has"][tag="VVN"]と現在完了形を意味する式を書きます。

右側の部分でdoes not containを選び同様に、[word="have|has"][tag="VVN"]という現在完了形を意味するCQLを書きます。

そうすると学習者の英文には現在完了があるが、その添削文には現在完了がない文、つまり学習者が間違っで現在完了を使っている文を抽出することができます。

The screenshot displays the ICNALE_EE search interface. The left panel shows the 'ADVANCED' search settings. The 'Query type' dropdown is set to 'CQL', and the query '[word="have|has"][tag="VVN"]' is entered in the CQL field. The right panel shows the search results for 'English, ICNALE_EE_EDIT'. The 'where' dropdown is set to 'does not contain', and the same CQL query '[word="have|has"][tag="VVN"]' is entered in the CQL field. Red circles highlight the 'does not contain' filter and the CQL query in both panels.