

Lingua Corpus Network

TUFS Lingua Corpus Network

検索用ソフト使用マニュアル

Guide Book for Corpus Retrieving Software

PERC Corpus 編

東京外国語大学ワールド・ランゲージ・センター

2024年9月

東京外国語大学ワールド・ランゲージ・センター

目 次

PERC コーパスについて	エラー! ブックマークが定義されていません。
1 ご利用の前に	4
(1) 短縮形の扱い	4
(2) n't の基本形は not	5
(3) 群前置詞、群副詞、外来語の扱い	5
(4) 単語認識について疑問が生じたら	5
(5) 本文中に現われる、& で始まる文字列について	7
文字エンティティリスト	7
2 検索の種類と切換	15
■ 「語（句）検索」と「コロケーション」、「単語情報」	15
(1) 「語（句）検索」	15
(2) 「コロケーション」	15
(3) 「単語情報」	15
3 「語（句）検索」	16
■ 検索の基本設定	16
(1) 「基本形で検索する」	16
(2) アステリスク記号を使った近傍検索	17
(3) 「2 単語間の語数」	17
(4) 「出現位置指定」	18
(5) 短縮形の入力方法	18
(6) 品詞を指定した検索	18
■ 品詞名（略名）による品詞指定	21
■ 品詞指定時のヒント	21
(1) 検索のシンタクス	22
(2) NOT 検索の指定について	23
(3) 検索例	23
「サブコーパス」の設定	24
4 KWIC 画面（語(句)検索の結果表示画面）	25
■ 結果の数値の見方	25
(1) 「総件数」	25
(2) 「採用」	26
(3) 検索結果画面のページ数	26
■ KWIC 画面から原文を表示する	26
■ KWIC 画面の表示変更	27
(1) 「初期設定」	27
(2) 「表示変更」	28

(3)	「結果表示」	29
(4)	「属性表示」	30
■ KWIC 画面のソート機能		30
(1)	ソートの設定	30
(2)	「指定なし」	31
(3)	「左ソート」	32
(4)	「右ソート」	32
(5)	「中心語を含む右ソート」、「中心語を含む左ソート」	33
(6)	「詳細ソート」	33
(7)	「ソートキー」の設定	35
(8)	「大文字/小文字を区別」	35
5	KWIC 画面のダウンロード	37
6	「クラスター」検索	37
■ 「集計範囲」		38
■ 「品 詞」		40
(1)	品詞一覧表の見方と利用方法	40
(2)	異なる品詞を複数指定する方法	41
■ 「頻度」、「集計キー」、「ソート」		41
■ 「ダウンロード」		41
7	「分布グラフ」	42
8	「コロケーション」	44
■ 中心語の指定		44
(1)	「基本形で検索」	44
(2)	中心語、共起語の品詞設定	44
■ 共起語の設定		45
■ 「共起語のソート」		45
(1)	「Tスコア」	45
(2)	「LogLogスコア」	45
(3)	「M.I.スコア」	45
■ 「共起表」の見方		46
(1)	出現頻度	47
(2)	「-3～-1」と「+1～+3」の見方と集計方法	48
(3)	共起表最上段の数値表示について	48
■ 表示設定		48
■ KWIC 画面表示		49
■ 「ダウンロード」		49
9	「単語情報」	49
■ 検索条件		50
■ KWIC 表示と原文表示		51
10	「ログアウト」	52

■ ログアウトの方法	52
■ 「ログアウト」ボタンを押さずに、ブラウザーを閉じた場合に起こること	52
■ なんらかの障害で正常にログアウトできなかった場合に起こること	53

著作権について

TUFS Lingua Corpus Network の著作権は東京外国語大学ワールド・ランゲージ・センターが保有しています。また、PERC Corpus の著作権は Professional English Research Consortium が保有しています。検索ソフトの著作権は東京外国語大学ワールド・ランゲージ・センターが保有しています。

<個人的に検索結果を使用される場合>

研究・教育などの個人使用の範囲において TUFS Lingua Corpus Network のコーパス検索結果を引用される場合は、必ずウィンドウ下部に表示されている著作権情報を明示してください。
コーパスに収録されているテキストはすべて原著者が著作権を所有しており、日本国著作権法によって保護されております。KWIC 以外の原文表示は、公正かつ適正な引用の範囲を超える場合は、原著者の許諾をお取りください。

<検索結果を商用利用される場合>

書籍などの商用出版物（デジタルを含む）に、TUFS Lingua Corpus Network のコーパス検索結果を引用される場合は、必ずウィンドウ下部に表示されている著作権情報を明示してください。なお、検索画面などを掲載される場合は、事前に東京外国語大学ワールド・ランゲージ・センターの許諾をお取りください。

コーパスに収録されているテキストはすべて原著者が著作権を所有しており、日本国著作権法によって保護されております。KWIC 以外の原文表示は、公正かつ適正な引用の範囲を超える場合は、原著者の許諾をお取りください。

1 ご利用の前に

(1) 短縮形の扱い

否定の n't や所有を表す 's は一語として扱われます。従って、これらを含む連語を検索する場合には、必ず直前の語を空白で区切ってください。

例：don't stop を調べたい場合は、do n't stop と入力します。

例：God's sake を調べたい場合は、god 's sake と入力します。

このほかの以下のようないくつかの短縮形も、n't や's と同様必ず直前の語とは空白で区切ってください。

'll 'd 'm 're 've na (wanna→wan na, gonna→gon na)

(2) n't の基本形は not

n't は 1 語として扱われますが、n't の lemma 形 (Lingua Corpus Network では「基本形」と呼びます) は not です。(詳しくは後述する「3 「語句検索」」の「■検索の基本設定」をご覧ください) 活用をすべて含めて検索したい場合は「基本形」で検索しますので、

例：isn't や wasn't など be 動詞の活用を無視して否定形を全て調べたい場合は、be not と入力します。

同様に、上でふれたその他の短縮形の「基本形」は次のようにになります。

'll → will	'd → would, had'm → be	're → be	've → have
na → to	's → 's (所有), be, have		

wan na の wan の基本形は want で、gon na の gon の基本形は go なので、wan na の基本形は want to となり、gon na の基本形は go to となります。

(3) 群前置詞、群副詞、外来語の扱い

in spite of や on behalf of のようなイディオムは一語として扱われ、単一の品詞コードが振られています。このような語句を検索する場合には、空白を半角の下線に置き換えて一語として検索してください。

例：in spite of を調べたい場合は、in_spite_of と入力します。

例：in order を調べたい場合は、in_order と入力します。

例：a la mode を調べたい場合は、a_la_mode と入力します。アクセント記号は無視します。

as soon as のような接続詞や、in order のような副詞も一語として扱われます。また、a la mode のような外来語も同様です。

例：as soon as を調べたい場合は、as_soon_as と入力します。

(4) 単語認識について疑問が生じたら

BNC の公式サイトに、群前置詞や群副詞のリストがありますので、もし検索結果が非常に少ない場合や、疑問が生じた場合には確認してください。a bit のように全体で副詞の場合と文字通り冠詞+名詞の二語の場合を文法機能上、区別している連語も少なからずあります。

以下の図は、公式サイト上の一覧リストの抜粋です。

左から (1) 連語 (ABC 順)、(2) 一語に扱われる場合の品詞コード、(3) 連語が分割されて一語づつ扱われる場合の品詞コード、(4) 注記となっています。自分が調べたい連語が (1) の欄にあれば、下線でつないで一語として検索してください。

(1) Multiword type	(2) Multiword tag, eg <w AV0>a bit	(3) Alternative tags when not a multiword, eg <w AT0>a <w NN1>bit	(4) Notes
number + 's	CRD	CRD + VBZ/VHZ/POS	Eg <w CRD>6's and <MC2>7's
letter + 's	ZZ0	ZZ0 + VBZ/VHZ/POS	Eg <w ZZ0>A's and <ZZ2>B's
a bit	AV0	AT0 + NN1	<u>1</u>
a capella	AJ0 / AV0		
a fortiori	AV0 / AJ0		
a good deal	AV0	AT0 + AJ0 + NN1	<u>1</u>

マルチワードに関しての BNC 公式サイト : <http://www.natcorp.ox.ac.uk/docs/multiwd.htm>

なお、「単語情報」でこれらの連語を調べることもできますので、活用してください。以下は「単語情報」から as_soon_as を引いた例で、全体が接続詞一語として扱われます。

No.	基本形 □ ▾	基本形合計 □ ▾	表記形 □ ▾	表記形合計 □ ▾	品詞頻度
1	as_soon_as	96	as_soon_as	96	CJS 96

総件数 : 96件 (0.59 / 10万語) 採用 : 96件 基本形 サブコーパス指定: 無			
1	ackground , the gel was put into 7-5% (v/v) acetic acid	as_soon_as	the bands had developed [29] .
2		As_soon_as	a membrane-bound preBCR is formed the preB-II cell shuts
3	nasal and lens placodes stain precociously for N-cadherin	as_soon_as	they can be recognized as ectodermal thickenings .
4	ucose under hypoglycaemic conditions , reaching a maximum	as_soon_as	glucose becomes higher than the physiological level (bas
5	clamps , in which the glucose entry rate becomes minimal	as_soon_as	the plasma insulin concentration exceeds 0.35 nmol/l (De
6		As_soon_as	the carotid bifurcation had been appropriately prepared ,
7	ing film and bacterial deposition on urogenital substrata	As_soon_as	a device is inserted into the urogenital tract , it becom
8		As_soon_as	the timescales (i.e. velocities) of the electronic and
9		As_soon_as	we accept that individuals have different preferences , u
10	ic procedure one can tentatively stop the cooling process	as_soon_as	the temperature has reached a wanted level .

※ ただし、"as soon as possible"のように文が続かずに副詞句の場合は、三語として扱われます。

All samples were frozen as soon as possible after collection to reduce any degradation of 1080
 these supernatants were then stored at 4 degrees and analysed as soon as possible after filtration .
 ithin hours It is important to be able confirm the diagnosis as soon as possible and to instigate the correct treatment , particular
 uces the rate at which it sends , the others should increase as soon as possible .
 king on a link , for_instance , wants to see a page or image as soon as possible .
 areas of the project and having team members look into them as soon as possible .
 The techniques include as soon as possible (ASAP) , as late as possible (AMP) , list , fre

(5) 本文中に現われる、& で始まる文字列について

KWIC 画面の本文に、"や"のような意味不明の文字が現われることがあります。これは、アスキイの文字コードセットにない外字を表すために利用される標準的な記法で「文字実体」と呼ばれます。例えば、"は「開始の引用符」、"は「終了の引用符」を表します。ブラウザ側が解釈して適当なフォントを割当てた場合には、KWIC 上でも適切に表示されますが、大抵の場合解釈されません。なお、この PERC Corpus では、数学、物理、化学などの数式や公式、方程式、さらに一連の記号やマークなどが続く部分は、すべて"で置き換えられています。あらかじめご諒解ください。

KWIC 画面上に現われる文字実体の内容が知りたい場合には、大変お手数ですが下記の文字エンティティリストを参考にしてください。

文字エンティティリスト

character	entity	description
Á	Á	capital A, acute accent
á	á	small a, acute accent
ă	ă	small A, breve
Â	Â	capital A, circumflex accent

â	â	small a, circumflex accent
'	´	acute accent
Æ	Æ	capital AE diphthong (ligature)
æ	æ	small ae diphthong (ligature)
&agr;	&agr;	small alpha, Greek
À	À	capital A, grave accent
à	à	small a, grave accent
Ā	Ā	capital A, macron
ā	ā	small a, macron
&	&	ampersand
ą	ą	small a, ogonek
≈	≊	approximate, equals
Å	Å	capital A, ring
å	å	small a, ring
*	*	asterisk
ã	ã	small a, tilde
Ä	Ä	capital A, dieresis or umlaut mark
ä	ä	small a, dieresis or umlaut mark
&Bgr;	&Bgr;	capital Beta, Greek
&bgr;	&bgr;	small beta, Greek
&bquo;	&bquo;	normalized begin quote mark
¥	\	reverse solidus
•	•	round bullet, filled
ć	ć	small c, acute accent
Č	Č	capital C, caron
č	č	small c, caron
Ҫ	Ç	capital C, cedilla
ҫ	ç	small c, cedilla
߂	ĉ	small c, circumflex accent
߄	¢	cent sign
߁	✓	tick, check mark
߃	○	circle, open

^	ˆ	circumflex accent
@	@	commercial at
©	©	copyright sign
↓	↓	downward arrow
-	‐	hyphen (true graphic)
d'	ď	small d, caron
°	°	degree sign
&Dgr;	&Dgr;	capital Delta, Greek
&dgr;	&dgr;	small delta, Greek
..	¨	dieresis
÷	÷	divide sign
\$	$	dollar sign
đ	&dstroke;	small d, stroke
▼	▾	dn tri, filled
É	É	capital E, acute accent
é	é	small e, acute accent
Ě	Ě	capital E, caron
ě	ě	small e, caron
Ê	Ê	capital E, circumflex accent
ê	ê	small e, circumflex accent
&eegr;	&eegr;	small eta, Greek
&Egr;	&Egr;	capital Epsilon, Greek
&egr;	&egr;	small epsilon, Greek
È	È	capital E, grave accent
è	è	small e, grave accent
ē	ē	small e, macron
ę	ę	small e, ogonek
=	=	equals sign
&equo;	&equo;	normalized end quote mark
ð	ð	small eth, Icelandic
Ё	Ë	capital E, dieresis or umlaut mark
ë	ë	small e, dieresis or umlaut mark

\flat	♭	musical flat
$\&formula;$	&formula;	mathematical formula
$\frac{1}{2}$	½	fraction one-half
$\frac{1}{3}$	⅓	fraction one-third
$\frac{1}{4}$	¼	fraction one-quarter
$\frac{1}{5}$	⅕	fraction one-fifth
$\frac{1}{6}$	⅙	fraction one-sixth
$\frac{1}{7}$	&frac17;	fraction one-seventh
$\frac{1}{8}$	⅛	fraction one-eighth
$\frac{1}{9}$	&frac19;	fraction one-ninth
$\frac{2}{3}$	⅔	fraction two-thirds
$\frac{2}{5}$	⅖	fraction two-fifths
$\frac{3}{4}$	¾	fraction three-quarters
$\frac{3}{5}$	⅗	fraction three-fifths
$\frac{3}{8}$	⅜	fraction three-eighths
$\frac{4}{5}$	⅘	fraction four-fifths
$\frac{4}{7}$	&frac47;	fraction four-sevenths
$\frac{5}{6}$	⅚	fraction five-sixths
$\frac{5}{8}$	⅝	fraction five-eighths
$\frac{7}{8}$	⅞	fraction seven-eighths
$\&ft;$	&ft;	feet indicator
\geq	≥	greater-than-or-equal
$\&Ggr;$	&Ggr;	capital Gamma, Greek
$\&ggr;$	&ggr;	small gamma, Greek
$\`$	`	grave accent
\gg	≫	dbl greater-than sign
$\>$	>	greater-than sign
$\frac{1}{2}$	½	fraction one-half
\heartsuit	♥	heart suit symbol
\cdots	…	ellipsis (horizontal)
$\overline{}$	―	horizontal bar
\hbar	ħ	small h, stroke

Í	Í	capital I, acute accent
í	í	small i, acute accent
Î	Î	capital I, circumflex accent
î	î	small i, circumflex accent
¡	¡	inverted exclamation mark
&igr;	&igr;	small iota, Greek
ì	ì	small i, grave accent
Ï	ī	small i, macron
∞	∞	infinity
&ins;	&ins;	inches indicator
¿	¿	inverted question mark
Ï	Ï	capital I, dieresis or umlaut mark
ï	ï	small i, dieresis or umlaut mark
&kgr;	&kgr;	small kappa, Greek
&khgr;	&khgr;	small chi, Greek
Ĺ	Ĺ	capital L, acute accent
ĺ	ĺ	small l, acute accent
←	←	leftward arrow
{	&lcurb;	left curly bracket
≤	≤	less-than-or-equal
&lgr;	&lgr;	small lambda, Greek
_	_	low line
[[left square bracket
Ł	Ł	capital L, stroke
ł	ł	small l, stroke
«	≪	double less-than sign
<	<	less-than sign
—	—	em dash
&Mgr;	&Mgr;	capital Mu, Greek
&mgr;	&mgr;	small mu, Greek
µ	µ	micro sign
·	·	middle dot

ń	ń	small n, acute accent
ჳ	♮	music natural
њ	ň	small n, caron
ɳ	ņ	small n, cedilla
-	–	en dash
&ngr;	&ngr;	small nu, Greek
Ñ	Ñ	capital N, tilde
ñ	ñ	small n, tilde
#	#	number sign
Ó	Ó	capital O, acute accent
ó	ó	small o, acute accent
Ô	Ô	capital O, circumflex accent
ô	ô	small o, circumflex accent
Œ	Œ	capital OE ligature
œ	œ	small oe ligature
&Ogr;	&Ogr;	capital Omicron, Greek
&ogr;	&ogr;	small omicron, Greek
ò	ò	small o, grave accent
&OHgr;	&OHgr;	capital Omega, Greek
&ohgr;	&ohgr;	small omega, Greek
Ω	Ω	ohm sign
ō	ō	small o, macron
Ø	Ø	capital O, slash
ø	ø	small o, slash
Õ	Õ	capital O, tilde
õ	õ	small o, tilde
Ö	Ö	capital O, dieresis or umlaut mark
ö	ö	small o, dieresis or umlaut mark
%	%	percent sign
&Pgr;	&Pgr;	capital Pi, Greek
&pgr;	&pgr;	small pi, Greek
&PHgr;	&PHgr;	capital Phi, Greek

&phgr;	&phgr;	small phi, Greek
+	+	plus sign
±	±	plus-or-minus sign
£	£	pound sign
"	″	double prime or second
'	′	prime or minute
&PSgr;	&PSgr;	capital Psi, Greek
&psgr;	&psgr;	small psi, Greek
"	"	quotation mark
ŕ	ŕ	small r, acute accent
√	√	surd = radical (square root)
→	→	rightward arrow
Ř	Ř	capital R, caron
ř	ř	small r, caron
}	}	right curly bracket
®	®	registered sign
&rehy;	&rehy;	maps to soft hyphen
&rgr;	&rgr;	small rho, Greek
]]	right square bracket
Ś	Ś	capital S, acute accent
ś	ś	small s, acute accent
Š	Š	capital S, caron
š	š	small s, caron
§	Ş	capital S, cedilla
§	ş	small s, cedilla
ˆ	ŝ	small s, circumflex accent
§	§	section sign
&Sgr;	&Sgr;	capital Sigma, Greek
&sgr;	&sgr;	small sigma, Greek
#	♯	musical sharp
&shilling;	&shilling;	British shilling
~	∼	similar

/	/	solidus
¹	¹	superscript one
²	²	superscript two
³	³	superscript three
ß	ß	small sharp s, German (sz ligature)
ť	ť	small t, caron
đ	ţ	small t, cedilla
&tgr;	&tgr;	small tau, Greek
&THgr;	&THgr;	capital Theta, Greek
&thgr;	&thgr;	small theta, Greek
þ	Þ	capital THORN, Icelandic
þ	þ	small thorn, Icelandic
×	×	multiply sign
™	™	trade mark sign
Ú	Ú	capital U, acute accent
ú	ú	small u, acute accent
Û	Û	capital U, circumflex accent
û	û	small u, circumflex accent
&Ugr;	&Ugr;	capital Upsilon, Greek
&ugr;	&ugr;	small upsilon, Greek
ù	ù	small u, grave accent
ū	ū	small u, macron
„	¨	umlaut mark
ö	ů	small u, ring
Ü	Ü	capital U, dieresis or umlaut mark
ü	ü	small u, dieresis or umlaut mark
	|	vertical bar
ŵ	ŵ	small w, circumflex accent
&xgr;	&xgr;	small xi, Greek
ý	ý	small y, acute accent
Ŷ	Ŷ	capital Y, circumflex accent
ŷ	ŷ	small y, circumflex accent

¥	¥	yen sign
Ý	Ÿ	capital Y, dieresis or umlaut mark
ÿ	ÿ	small y, dieresis or umlaut mark
Ź	ź	small z, acute accent
Ž	Ž	capital Z, caron
ž	ž	small z, caron
ż	ż	small z, dot above
&Zgr;	&Zgr;	capital Zeta, Greek
&zgr;	&zgr;	small zeta, Greek

2 検索の種類と切換

■ 「語（句）検索」と「コロケーション」、「単語情報」

検索の種類は、「語（句）検索」、「コロケーション」、「単語情報」の3種類があります。ログイン直後の画面は、常に図のようになっており、「語（句）検索」を開始できる状態です。これを「検索画面」と呼びます。

他の検索に切り替える場合はこの3つのタブから、希望の検索をクリックします。



(1) 「語（句）検索」

検索語を含む文を検索するための機能です。検索結果は、KWIC画面に表示されます。用例を探したい通常の目的には、「語（句）検索」の利用が便利です。

(2) 「コロケーション」

ある語が文中でどのような語と一緒に使われるかについて、その語の前後5語の範囲にわたって出現する語の頻度を集計し、その結果を表で表示します。

(3) 「単語情報」

ある単語を指定して、その単語のBNCにおける語形別、および品詞別の出現頻度一覧を表示することができます。

3 「語（句）検索」

単語や連語で検索できます。連語で引く場合は、単語の間に半角の空白を必ず入れます。図は、検索語に"give up"を指定して入力エリアに入力した例です。



検索語に使用できる文字種は ASCII で定義されているアルファベットと数字です。ただし、大文字と小文字は区別されません。カンマ、ピリオド、感嘆符、疑問符などの記号も使用できます（一部の記号は検索できません）。ただし、記号も 1 語と扱われるため、記号と語との間には、次のように、必ずスペースを入れます。



検索語を入力したあと、「検索」ボタン、または Enter キー、実行キーを押すと検索が開始されます。

■ 検索の基本設定

語(句)検索は、単に語や語句で検索する以外に、目的に応じて以下の 3 つの基本的指定が可能です。これらの指定は組み合わせて使うことができます。

(1) 「基本形で検索する」

基本形は辞書の見出し語の意味 (lemma で、語形変化していない綴りです。活用や单数複数などの語形変化をすべて含めた検索をしたい場合に、チェックボックスをONにします。ONになっている場合には、検索語の入力エリアには、基本形しか指定することができませんのでご注意ください。

語形変化として扱われるのは、形容詞の比較級と最上級、名詞の单数形と複数形、動詞の活用形です。ただし、基本形は 100% 完全に校正されていませんので、「単語情報」で語形別の頻度や用例を確認されることをお奨めします。図は、"look"の活用をすべて含めて検索をする例です。look, looks, looking, looked"が検索されます。

この「基本形で検索」が選択されていない場合は、入力エリアに入力された形(「表記形」と呼びます)のままで検索されます。

(2) アステリック記号を使った近傍検索

離れた 2 語を検索したい場合には、検索したい 2 語の間に、半角の* (アステリック) を入れます。ただし、必ず*の前後にスペースを入れてください。例えば、検索語が look と closely なら、"look * closely"と指定します。*は 1 回しか使えません。なお、"look closely"のように連続している場合も、この検索結果には含まれます。また、"not only * but also"のように、前後は連語でもかまいません。

"*"の部分に相当するのは、単語（連語）であり、文字ではありません。したがって、綴りが思い出せない単語の検索や、動詞の語形変化を無視したような曖昧検索に利用することはできません。"look *"という検索語から、"looks, looking, looked"の例を引くことはできません。

(3) 「2 単語間の語数」

2 単語の間に挟まれる語数の初期設定は「1 語以内」になっています。これを最大 9 語以内まで増やすことができます。なお、文の区切りを超えた検索は行いません。

1語以内
2語以内
3語以内
4語以内
5語以内
6語以内
7語以内
8語以内
9語以内

(4) 「出現位置指定」

検索語が出現する位置を、「文頭」か「文末」に指定できます。V印をクリックするとリストが出ます。図は、否定の命令を調べるために、"do not"を検索語にして、「文頭」に出現する例だけを検索する例です。「文末」を選んだ場合には、必ず適切な文末記号（ピリオドや疑問符）まで指定します。



(5) 短縮形の入力方法

短縮形は、独立した1語として扱われます。また所有の'sも1語です。下記の例を参考にしてください。必ずスペースで区切って入力してください。

don't → do n't wanna → wan na Napoleon's → Napoleon 's

(6) 品詞を指定した検索

検索語(句)の各単語に品詞を指定して検索する場合は「品詞指定」ボタンから品詞を選択して検索語句に品詞条件を追加することができます。



操作手順：

- i) 検索したい語(句)を入力します。その際、表記と異なる変化形を含む品詞を追加する場合は「基本形で検索」をチェックします。(たとえば複数形や過去形、比較級なども取得したい場合)
- ii) 「品詞指定」ボタンをクリックして品詞表を表示させます。
- iii) 検索窓に入力されている場合には検索窓にある各ユニットが順に選択欄に表示されますので、品詞を指定する単語を選択します。それに対する品詞を選択し、最後に「検索窓にセットする」ボタンをクリックして設定を反映させます。

検索語は左から順に①の単語、②の単語、...となります。

例：star * with の場合

"*"を使う場合は「2 単語間の語数」の指定がいくつであっても1語として数えます。

検索窓に入力された語句が選択欄に順に表示されるので任意の品詞の組合せを選択します。

例

検索窓に入力された語句が選択欄に順に表示されるので任意の品詞の組合せを選択します。

選択してください

- ① [star]
- ② [*]
- ③ [with]

検索窓にセットする

例

検索窓に入力された語句が選択欄に順に表示されるので任意の品詞の組合せを選択します。

選択してください

- ① [star]
- ② [*]
- ③ [with]

検索窓にセットする

例：star~P(VV.*)* with (すでにstarに一般動詞が指定されている品詞を普通名詞に変更する例)

検索語(句)の各単語に品詞条件を追加します：

選択してください

- ① [star~P(VV.*)]
- ② [*]
- ③ [with]

ツトする

例

検索語(句)の各単語に品詞条件を追加します：

選択してください

- ① [star~P(VV.*)]
- ② [*]
- ③ [with]

ツトする

既に品詞が指定されている単語に対して、品詞指定を行う場合はそれまでの条件をクリアして新たに

指定した品詞と入れ替えになります。

検索例：

- i) 次の例は star * with の star に一般動詞を指定した例です。特に検索語句に多品詞の単語が含まれる場合は、品詞条件を付与することで、より絞り込んだ検索結果が期待できます。

品詞を指定した検索は、**単語** と **～(品詞コード)** を組合せて指定します。

* 品詞表からの追加を使わないので手入力でも同じ結果となります。

その際、**～()** は "～" (チルダ)や括弧を含め**全て半角**で括弧内はスペースは入れられません。また複数の品詞を指定する場合は、下の例のように" | " (半角) を使います。

この時、一般動詞は **VV.** 動詞全般は **V..** などと "." ドットを使って省略することもできます。同様に名詞も **N..** 等と入力することで同じ条件を指定したことになります。

品詞条件は、指定した品詞を含む場合と含まない場合を選択できます。

- ii) 品詞条件は、指定した品詞を含む場合と含まない場合を選択できます。

通常指定した品詞を含む場合は **～(N..)** と指定します。 **～～(N..)** と "～" (チルダ)を2つ続けることで指定した品詞を含まない条件での検索を行うことになります。

- iii) 品詞条件は複数の単語にそれぞれ付与することもできます。

品詞表から順次選択して追加していきます。また、品詞表を使って、一度指定した単語の品詞条件を変えて再指定する場合は検索語(句)入力欄にすでにある品詞と入れ替わります。

制限事項： **2** 単語間の語数を指定する "*" に品詞を指定することはできません。

■ 品詞名（略名）による品詞指定

品詞指定は品詞表に見られるように PERC Corpus で利用されている C5 品詞記号だけでなく、よりわかりやすく入力できるように、辞書などで利用される略名による指定を代表的な品詞について指定できるようにしました。

	品詞名による指定	ワイルドカード指定
名詞	noun	N. *
代名詞	pron	PN. * DT0
形容詞一般	adj	AJ. *
副詞、副詞辞	adv	
接続詞	conj	CJ. *
前置詞	prep	
動詞	verb	VB. VD. VH. VV.
動詞〔過去〕		V. D
動詞〔-ing 形〕		V. G
動詞〔原形不定詞〕		V. I
動詞〔過去分詞〕		V. N
動詞〔3人称単数〕		V. Z

※ V.* は動詞を全て指定できますが、さらに VMO (can や will などの法助動詞) も含まれてしまい、ますので、全ての動詞のみを指定する場合には "verb" または "VB. |VD. |VH. |VV." を使います。

■ 品詞指定時のヒント

品詞表の左側の大分類、特に「名詞」「動詞」(be 動詞、do 動詞などの第2分類なども含む) を指定をした場合は検索語との組合せにご注意ください。

例： test~(NN0|NN1|NN2)

この例のような指定をした場合、「基本形で検索」のチェックがない場合には test (名詞) の形) 表記系) のみの検索になり、複数形 "tests" の用例は検索できませんので、複数形の用例も得たい場合には「基本形で検索」にチェックを入れます test は名詞の単数形を表す NN1 なので、名詞の複数形を表す NN2 を指定しても該当するものがなく、tests はヒットしません)。

一方で、入力した単語が基本形 (辞書の見出し語相当の形) とは別の複数形や変化形をそのまま入力している場合は、『基本形で検索』にチェックしてしまうと、検索できない場合がありますので注意が必要です。

例： tests~(NN2|VVZ) ⇒

入力した単語は「基本形」ではないので『基本形で検索』にチェックしてしまうと、検索できません

特に、他の文書からフレーズをコピーしたものを検索語句にペーストした際は、多くの場合、変化形などをそのまま含んでいますので、品詞を指定する際はご注意ください。

一般的に変化形なども含め、用例を広く収集したい場合には、検索語(句)は見出し語化した単語（基本形）に直して入力し、『基本形で検索』にチェックします。その上で品詞を指定する場合には "動詞" や "名詞" など指定範囲が大きくなるように品詞を選択します。

■ 品詞指定を用いた詳細検索

前述した検索の基本設定 「(6) 品詞を指定した検索」 を拡張して、具体的な単語を指定せず単に品詞のみを指定したフレーズも検索することができます。また、品詞と同様に単語についても複数列記して OR 検索もご利用いただけます。

(1) 検索のシンタクス

検索語入力欄にスペースで区切られたユニットが 1 つの「単語」に対応します。

詳細検索を指定する際には 1 ユニットが $\sim W()$ など チルダ " \sim " と 半角の括弧 () で構成されています。

1 ユニットの構成要素 $\sim L()$, $\sim W()$, $\sim P()$

チルダの直後につづく大文字 L,W,P(のいずれか 1 文字)はその後に続く括弧内の内容がどの属性を表しているかを示しています (L=lemma、W=word、P=part of speech)。

括弧内には $\sim W$ であれば「表記形」を、 $\sim L$ であれば「基本形」を、 $\sim P$ であれば「品詞」を入力します。それぞれ、 $\sim W$ (表記形)、 $\sim L$ (基本形)、 $\sim P$ (品詞) となります。また、OR 検索には"|" (半角) を用います。

	1 ユニットの記述の具体例
i) 表記形の指定	$\sim W(apple)$
ii) 表記形と品詞の組合せ	$\sim W(apple)\sim(N.*)$ 、又は $\sim W(apple)\sim P(N.*)$
iii) 基本形の指定	$\sim L(apple orange)$
iv) 基本形と品詞の組合せ	$\sim L(center centre)\sim(N.*)$ 、又は $\sim L(center centre)\sim P(N.*)$
v) 品詞の指定 (一般動詞)	$\sim P(VV.*)$ (※ be, have, do は含まれない)
vi) 品詞の指定 (動詞全般)	$\sim P(verb)$

上の例のうち、 $\sim W(apple|orange)$ や $\sim (apple|orange)$ のように具体的な単語を検索する場合は、次のように W や L を省略することができます。

$\sim (apple|orange)$ (※ チルダ " \sim " は省略できません)

その際は「基本形で検索」にチェックがしてあれば $\sim L(apple|orange)$ と解釈されることなり、チェックがなければ $\sim W(apple|orange)$ と解釈されることになります。また、「基本形で検索」のチェックがない場合（以後「表記形の検索モード」とします）で任意の単語のみ基本形にしたい場合は、たとえば $\sim L(apple)$ とすれば、指定した単語（ここでは apple）のみ基本形を指定して検索することになります。

ただし品詞のみで単語を指定する場合には $\sim P(NN1|NN2)$ 等と P を省略することはできません。また "|" を使わない場合は $apple$ と単に指定すれば通常の検索語入力になります。

これまでの検索語の単語の指定方法 $apple pie$ は「表記形の検索モード」では $\sim W(apple) \sim W(pie)$ と同じになります。「基本形の検索モード」では $\sim L(apple) \sim L(pie)$ と同じになります。 $\sim L, \sim P, \sim W$ を入れたほうが誤って検索することがおこりにくくなりますので、慣れるまでは省略しない書き方をおすすめします。

(2) NOT 検索の指定について

NOT 検索は、現在のところ品詞の指定時のみ有効になります。また、制約事項として、検索語(句)の最初の単語を品詞のみ指定する、また "*" の直後に品詞のみ指定する場合も NOT を使用することができません。

検索例

- $\sim W(book | input) \sim \sim P(V.*)$ on
- $\sim W(round | bound) \sim \sim P(prep) \sim P(noun)$
- ✗ $\sim \sim P(V.*)$
- ✗ $\sim \sim W(pink | blue) colored \sim P(N.*)$
- ✗ $\sim W(get | have) * \sim \sim P(noun)$

(3) 検索例

検索したい語(句)の例	検索語句入力欄
形容詞 + examination	$\sim P(AJ.*)$ examination
品詞検索(一般動詞)	$\sim P(VV.*)$
core + 名詞	core $\sim P(N.*)$
make it * to 不定詞	$\sim L(make) it * to \sim P(VVI)$
find/give を使った二重目的語構文	$\sim L(find give) \sim P(PN.* DT0) \sim P(AT0) \sim P(N.*)$

※ 『基本形で検索』のチェックはいはずれもなしの場合

「サブコーパス」の設定

PERC Corpus のサブコーパスは、22 分野に分割されています。下図の「サブコーパス画面」から、22 サブコーパスを任意の組み合わせで指定することができます。指定の方法は、希望するサブコーパスの分類名やチェックボックスを直接選んで選択します。下図は、初期の設定で全てが選ばれている状態を示しています。

上部にある「全てを解除」を押すと一旦全てサブコーパスの指定が解除されますから、次いで希望するサブコーパスを選んで指定することもできます。指定が終了したら、必ず右上の「設定を保存する」ボタンを押します。指定したサブコーパスの語数と比率は、下段に表示されます。

サブコーパス設定

設定を保存する

| 全てを選択 | 全てを解除 | 選択を反転 |

サブコーパス名	ファイル数	語数	比率
<input checked="" type="checkbox"/> Agriculture	193	663,695	4.09
<input checked="" type="checkbox"/> Biology	399	1,410,344	8.70
<input checked="" type="checkbox"/> Chemistry	520	1,544,478	9.52
<input checked="" type="checkbox"/> Civil Engineering	110	403,628	2.49
<input checked="" type="checkbox"/> Computer Science	403	1,469,298	9.06
<input checked="" type="checkbox"/> Construction & Building Technology	13	44,068	0.27
<input checked="" type="checkbox"/> Earth Science	318	1,163,157	7.17
<input checked="" type="checkbox"/> Electrical & Electronic Engineering	172	632,179	3.90
<input checked="" type="checkbox"/> Engineering	518	1,929,699	11.90
<input checked="" type="checkbox"/> Environmental Sciences	10	47,991	0.30
<input checked="" type="checkbox"/> Fisheries	69	215,920	1.33
<input checked="" type="checkbox"/> Food Science	93	392,453	2.42
<input checked="" type="checkbox"/> Forestry	20	58,760	0.36
<input checked="" type="checkbox"/> General Science	68	289,761	1.79
<input checked="" type="checkbox"/> Materials Science	114	454,046	2.80
<input checked="" type="checkbox"/> Mathematics	67	310,834	1.92
<input checked="" type="checkbox"/> Medicine	1,125	3,330,238	20.54
<input checked="" type="checkbox"/> Metallurgy & Metallurgical Engineering	29	114,728	0.71
<input checked="" type="checkbox"/> Nuclear Science & Technology	46	159,271	0.98
<input checked="" type="checkbox"/> Oceanography	64	237,466	1.46
<input checked="" type="checkbox"/> Physics	325	1,229,505	7.58
<input checked="" type="checkbox"/> Telecommunications	23	114,506	0.71

語数 : 16,216,025 語 比率 : 100 %

4 KWIC 画面（語(句)検索の結果表示画面）

語(句)検索の結果は、デフォルトでは下の図のように KWIC 表示で示されます。このほかに、サブコーパス別の出現比率を視覚的にグラフで表す機能があります。(グラフ表示機能については「8 分布グラフ」を参照してください)

The diagram illustrates the process of displaying search results. It starts with a search result summary at the top, which includes a link to 'KWIC 表示' (KWIC display). Three arrows point down to the KWIC display interface below. The first arrow points to the 'KWIC 表示' link. The second arrow points to the 'KWIC' section of the display, which shows a list of search results. The third arrow points to the right side of the display, where a message about navigating between pages is shown.

サブコーパス間の出現頻度比率
をグラフで視覚的に表示しま
す。(「9 分布グラフ」を参照)

各行の特定の範囲を指定し、同じ形で出
現する語(句)(クラスター)を集計して表示
します。

検索された総件数。同じ文中
に出現する検索語(句)もカウ
ントされます。

KWIC の行を指定条件でソートし
て並べ替えます。

数字は、「現在のページ／総ページ」を示していま
す。<をクリックすると前のページに戻ります。>を
クリックすると、次のページに進みます。

語(句)検索結果 分布グラフ ON OFF ソート・表示変更 クラスター ダウン 検索語(句)をクリックすると原文と書誌情報を表示します

総件数 : 3889件 (23.98 / 10万語) 採用 : 3000件 表記形 サブコーパス指定: 無

1 a; to &formula; , where Ip is an inertia subgroup at p (see line 3 of p. 125 of [Flach 90]).
2 , the graph of C in the same region looks like a horizontal line .
3 The first error they make is on the fifth line of the r = 2 case on page 151 : The quoted work of Atkin and
4) and (3) in Section 2.1 or to a group having an invariant line of the complex , defined over &formula; .
5 by linear equations with real coefficients is called a real line in Pn C. Figure 6(a) shows the compactified amoeba of a real
6 in Pn C. Figure 6(a) shows the compactified amoeba of a real line in &formula; .
7 d with a three-dimensional surface plot in Maple , where the line &formula; is considered as a two-dimensional affine subspace
8 Let &formula; , and let &formula; be a real line in Pn C with Plucker coordinate &formula; .
9 eba directly in_terms_of the defining inequalities of a real line &formula; in 3- or 2-space .
10 Let &formula; be a line in P2 C given as the solution of the linear equation &formula;
11 d , n , the dimension of I , and the degree of I , the last line prints the degrees of a minimal set of generators of I. The
12 territory boundaries and connected by thick transparent fishing line (Siglon monofilament fishing line , Sunline Ltd , Kaisha ,
13 thick transparent fishing line (Siglon monofilament fishing line , Sunline Ltd , Kaisha , Japan) .
14 The line was positioned on the pole such_that the models would 'fly'

■ 結果の数値の見方

総件数 3889件 (23.98 / 10万語) 採用 3000件 表記形 サブコーパス指定: 無

(1) 「総件数」

PERC Corpus に含まれる件数を表します。括弧内の数字は、10 万語あたりの件数に換算した値です。総数にバラツキのあるサブコーパスの件数を比較する場合などは、この換算値を使って比較します。

(2) 「採用」

画面表示のためにサーバーが検索した KWIC 行の総数を表します。総件数のほうが大きい場合は、採用した件数は、後で説明する「初期設定」の「検索結果上限」に等しく、それ以下の場合は、総件数と同じです。

(3) 検索結果画面のページ数

The screenshot shows a search results page with various filters at the top: '分布グラフ ON OFF', 'ソート・表示変更', 'クラスター', and 'ダウンロード'. Below these are search parameters: '総件数: 3889件 (23.98 / 10万語) 採用: 3000件 表記形 サブコーパス指定: 無'. The main content area displays a single result line: '1 rmula; to &formula;, where Ip is an inertia subgroup at p (see line 3 of p. 125 of [Flach 90]).'. To the right of the content is a page navigation bar with buttons for '<', '1' (highlighted with a red oval), '/ 120', and '>'.

数字は、「現在のページ／総ページ」を示しています。< をクリックすると前のページに戻ります。

> をクリックすると、次のページに進みます。

Enter キー、実行キーを押すと、指定したページへジャンプします。

■ KWIC 画面から原文を表示する

色文字で表示されている KWIC の各行の検索語をクリックすると、別の画面が開き、前後の文脈をより広い範囲にわたって読むことができます。検索語は文中で赤の太字になっています。また、書誌情報が上段に表示されますので、出典を確認できます。

PERC Corpus	
File	00006
Category	Mathematics
Sub Category	MATHEMATICS
Title	Computing the Modular Degree of an Elliptic Curve
Author	Mark Watkins
Nation	USA
Journal title	EXPERIMENTAL MATHEMATICS
Volume	11 (Number 4 , Page 487-502)
Date Issue	2002
Publisher's name	A K Peters, Ltd.
ISSN	1058-6458

...orks in the discriminant field $Q_p(\#x2889AE28886;) = Q_p(\#x2889AE28892; 3) = Q_p(\#x2889A; 3)$.

Thus $H \varphi = C_4$, and $G \varphi = C_8$ is abelian.

This proves the theorem.

We next discuss $p = 2$.

Here the minimal twist will have neither 16 nor 64 exactly dividing its conductor (this follows from the table on page 121 of Coates-Schmidt, or more simply from an analysis of Tate's algorithm), so that δ_2 is neither 4 nor 6.

If δ_2 is odd, the Coates-Schmidt table tells us that $U_2(s) \equiv 1$ and $\delta_2 = (1 + \delta_2)/2$.

If $\delta_2 = 2$, again there is not much problem; the table says that $U_2(s) = (1+2/2s) \#x28892; 1$ and $\delta_2 = 1$.

The case of $\delta_2 = 8$ is the most difficult.

The appendix of Coates-Schmidt makes two errors, leading to the classification being incomplete.

The first error they make is on the fifth line of the $r = 2$ case on page 151: The quoted work of Atkin and Li requires the underlying form to have 16 dividing the level, and if the level of the absolute minimal quadratic twist of the form f associated to F (no longer necessarily rational, i.e., the twisted form can have a nontrivial Nebentypus character) is exactly divisible by 23, this does not apply.

An explicit example is the curve 768H (given by [0, 1, 0, 1, #x28892; 3]) for which the absolute minimal twist is of level 24.

Another error is on page 153 in the analysis of the case where the inertia subgroup is Q_8 , where they state that "G is obviously a 2-Sylow group of $GL_2(F_3)$, hence dihedral of order 16."

This should be semidihedral of order 16.

This causes them to miss the possibility that the absolute minimal twist can have 27 exactly dividing its level.

An example is 256B (given...).

■ KWIC 画面の表示変更



The screenshot shows the KWIC search interface. At the top right, there is a '初期設定' (Initial Settings) button with a red circle around it. Below the search bar, there is a 'ソート・表示変更' (Sort/Display Change) button with a red circle around it. The search bar contains the word 'line'. On the right side of the interface, there are dropdown menus for '出現位置指定' (Occurrence position specification) and '2単語間の語数' (Number of words between 2 words), both set to '指定なし' (Not specified). There are also buttons for 'ON' and 'OFF' under '分布グラフ' (Distribution graph), and a 'ダウンロード' (Download) button.

(1) 「初期設定」

KWIC 表示画面右上にある「初期設定」をクリックすると「検索結果の上限」と「KWIC の 1 ページ行数」を指定することができます。下の画面は、その設定画面です。デフォルトでは、検索結果の上限 3000 件、1 ページの行数 25 行となっています。

「検索結果上限」は、ヒット数のうち何件の KWIC データをサーバーからあなたの PC に送るかを決める数です。例えば、PERC コーパスに出現する the の総件数は、約百万件です。この全ての用例を表示することは、実際上不可能であり、また意味がありません。逆に少なければ、検索結果の信頼性を損ねる場合があります。できるだけ多くの検索結果を表示したい場合には、「検索結果上限」を大きな値にします。しかし、その分、サーバー側での処理に時間がかかります。サーバー側の処理が終わってから、最初のページが表示されますから、この値を大きく設定すると、最初の画面が表示されるまでの時間が長くなります。もし、処理を速くしたい場合は、「検索結果上限」を小さな値にします。

「検索結果上限」をたとえ小さく設定しても、検索結果は BNC 全体からランダムに選ばれますから、片寄った結果にはなりません。「検索結果上限」は、通常は読んでも苦にならない程度の行数に設定すれば、必要にして十分です。



(2) 「表示変更」

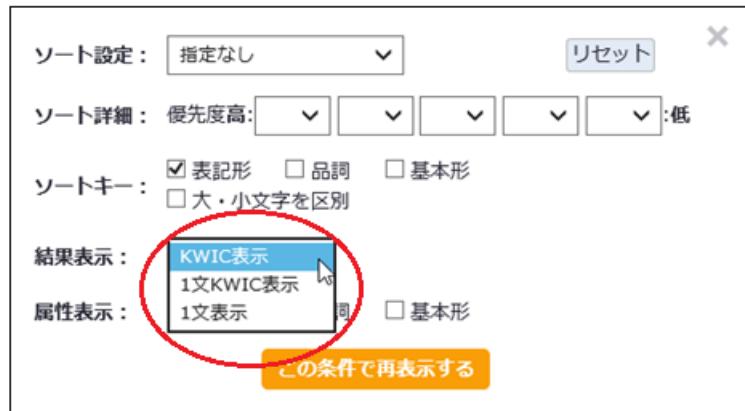
「ソート・表示変更」をクリックすると、下のようなウィンドウが開きます。

「結果表示」「属性表示」を使い検索結果の表示を変更します。

ソート設定 :	指定なし	リセット						
ソート詳細 :	優先度高:	□	□	□	□	□	□	低
ソートキー :	<input checked="" type="checkbox"/> 表記形	<input type="checkbox"/> 品詞	<input type="checkbox"/> 基本形					
結果表示 :	KWIC表示							
属性表示 :	<input checked="" type="checkbox"/> 表記形	<input type="checkbox"/> 品詞	<input type="checkbox"/> 基本形					

(3) 「結果表示」

「結果表示」の初期設定は「KWIC 表示」になっています。その右にあるVの印をクリックすると、次のようなプルダウンメニューが現れますので、ここから表示方法を選び、「この条件で再表示する」ボタンをクリックします。



初期設定は、通常の「KWIC 表示」で、長さが異なる前後の文脈を1行に収めるために、文の途中で切って長さを揃えた形式になっています。1行の幅は120文字です。

「1文 KWIC 表示」を選ぶと、図のように文の冒頭や最後で切ることをせずに1文全体を表示します。左右幅に収まらない文は複数行で表示されます。

語(句)検索結果 分布グラフ ON OFF ソート・表示変更 クラスター ダウンロード ◀ 12 / 120 ▶

総件数：3889件 (23.98 / 10万語) 採用：3000件 表記形 サブコーパス指定：無

276	Axial intensity profiles Optical line emission spectroscopy is used to measure the axial intensity profiles of the excited neutral (Ar , Mo) and ion (Ar +i , Mo +i , i = 1 , 2 , 3) species in the plasma .
277	After noise reduction and normalization to their respective peak values , the line intensity profiles for Ar I , Ar II , Ar III , Mo I , and Mo II are presented in Fig. 2 .
278	The intensity decay is related to the decrease of Ar I (the available neutral line (288.1 nm) profile . population for ionization is decreasing) and to the second ionization of the Ar ions as shown by the Ar III
279	Ar line profiles indicate that apart_from the plasma region close to the surface (0 to -50 mm) the plasma column is not affected by the presence of the sample .
280	The distance from the sample surface to the line peak intensity is indicative of the excitation mean free path of the Mo atoms and provides a meaningful way to measure excitation rate coefficient (see Section 3.3) .

「1文表示」を選ぶと、図のように検索語を画面の中心に揃えた KWIC 表示ではなく、文の冒頭を揃えた1文全体が表示されます。この場合でも検索語の場所がわかるように太字で表示します。

- 276 Axial intensity profiles Optical **line** emission spectroscopy is used to measure the axial intensity profiles of the excited neutral (Ar , Mo) and ion (Ar +i , Mo + i , i = 1 , 2 , 3) species in the plasma .
- 277 After noise reduction and normalization to their respective peak values , the axial **line** intensity profiles for Ar I , Ar II , Ar III , Mo I , and Mo II are presented in Fig. 2 .
- 278 The intensity decay is related to the decrease of Ar I (the available neutral population for ionization is decreasing) and to the second ionization of the Ar ions as shown by the Ar III **line** (288.1 nm) profile .
- 279 Ar **line** profiles indicate that apart_from the plasma region close to the surface (0 to -50 mm) the plasma column is not affected by the presence of the sample .
- 280 The distance from the sample surface to the **line** peak intensity is indicative of the excitation mean free path of the Mo atoms and provides a meaningful way to measure excitation rate coefficient (see Section 3.3) .

(4) 「属性表示」

「属性表示」から「品詞」を選ぶと、地の文に品詞コードを振った表示になります。



- 276 Axial|AJ0 intensity|NN1 profiles|VVZ Optical|AJ0 **line|NN1** emission|NN1 spectroscopy|NN1 is|VBZ used|AJ0 to|PRP measure
- 277 S respective|AJ0 peak|NN1 values|NN2 ,|PUN the|AT0 axial|AJ0 **line|NN1** intensity|NN1 profiles|NN1 for|PRP Ar|ITJ I|PNP ,|PUN Ar|ITJ
- 278 |ITJ ions|NN2 as|CJS shown|VVN by|PRP the|AT0 Ar|ITJ III|CRD **line|NN1** (|PUL 288.1|CRD nm|UNC)|PUR profile|NN1 ,|PUN
- 279 Ar|ITJ **line|NN1** profiles|NN2 indicate|VVB that|CJT apart_from|PRP the|AT0 pl
- 280 e|NN1 from|PRP the|AT0 sample|NN1 surface|NN1 to|PRP the|AT0 **line|NN1** peak|NN1 intensity|NN1 is|VBZ indicative|AJ0 of|PRF the|AT0

■ KWIC 画面のソート機能

(1) ソートの設定

KWIC の行ソートを指定するときに使います。「この条件で再表示する」を押すと設定した条件が有効になり、画面が閉じます。コロケーションや単語情報など別の検索モードに変更しない限り設定は保持されます。

ソート設定： 指定なし リセット

ソート詳細： 優先度高: ▼ ▼ ▼ ▼ ▼ ;低

ソートキー： 表記形 品詞 基本形
 大・小文字を区別

結果表示： KWIC表示

属性表示： 表記形 品詞 基本形

この条件で再表示する

(2) 「指定なし」

KWICの行をソートしません。図は、ソートの設定が「なし」で、“ask”を「基本形」で引いた例です。検索語の左右の語は揃っていません。

It is a natural question to ask	the growth of the order of An in the cases of Table 1 and 2 .
In_general , for a prime p , we could ask	how often p divides the modular degree .
Hilbert 's problem asks	for a classification of the topological types of real algebra
When Charlie asked	me to write an essay delineating the differences between inna
Transplantation Where , I am often asked	, is the Danger in a transplant ?
When I ask	the clinicians what plans they have for these remaining patie
tumor vaccines less than impressive , it seems reasonable to ask	whether a self/ nonself paradigm that makes sense in the cont

(3) 「左ソート」



KWIC の行を、検索語の左側の語でソートします。ソートキーは、検索語の左隣の語から順に前方に向けて自動的に設定されます。検索語の直前の語句で KWIC 行を揃えたい場合に使用します。図は、ソートの設定を「左ソート」にしてから、"ask"を基本形で引いた例の冒頭です。

1	Asking	about matters of importance to these people in REGARD TO thei
2	asked	their subjects to compare each picture stimulus with the prev
3	itivity ; 4) the freedom to choose modulation formats [ASK	, FSK , or PSK (amplitude/frequency/phase shift keying) &rs
4	· Asking	for information (interaction technologies) .
5	,000 a month to get a 1.5-Mbps T1 connection , and now we 're asking	them to give that same T1 connection to the home for \$30 a mo
6	In their classic study , Bisiach and Luzzatti (1978) asked	patients to describe the familiar Piazza del Duomo in Milan f
7	(1985) asked	subjects to perform a range of motor acts (such AS tongue pr
8	(1991) ask	how close to the derived consensus sequence does a sequence h

以下の例のように左隣に同じ単語がくる場合はさらにその左隣の単語をソートしています。

101	In this test subjects are asked to judge which end of a 200mm horizontal line a smaller line
102	doing so we present a final experiment in which subjects are asked to make a different decision at test .
103	Members of a quality team are asked to respond to the instrument with RESPECT TO the following ei
104	ceive two or three doses of cocaine or placebo , and then are asked to choose which of the initial test doses they would like to
105	ast with their relative slowness and inaccuracy when they are asked deliberately to acknowledge all of their errors by making the
106	e case when a controller has already been designed and we are asked to verify its operation , or to determine the sets of initial

(4) 「右ソート」

KWIC の行を検索語の右側の語でソートします。ソートキーは、検索語の右隣の単語から順に後方に自動的に設定されます。検索語の直後の語句で KWIC 行を揃えたい場合に使用します。

(5) 「中心語を含む右ソート」、「中心語を含む左ソート」

「中心語を含む左ソート」または「右ソート」を行いたい場合にこれを用います。「中心語」という呼称は必ずしも正確な表現ではありませんが、検索語がフレーズの場合は1番目の単語を中心（詳細ソートでの「0」に相当します）にしてこの語も含めてソートするという意味で便宜的に使っています。基本形で検索した際、検索語の語形ごとに比較したい場合などに利用できます。

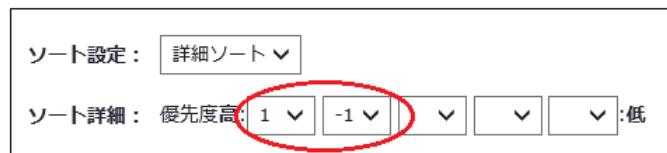
(6) 「詳細ソート」

「詳細ソート」では、ソートキーの場所と優先度を自由に設定できます。

「ソート設定」から「詳細ソート」を選んで、ソートを優先する順にソートキーを設定します。次の図は、プルダウンメニューを開いて第1キーを指定しているところです。「詳細ソート」の優先度を変更すると、自動的に「ソート設定」が「詳細ソート」に変更になります。



たとえば、検索語の右隣でソートして揃った行をさらに左隣でソートしたい場合には、(1, -1) の順に指定します。図は、直後の "for" で揃えて、次に検索語の直前で揃えた結果になっています。



IRI alerts her to that overhead and **asks** for confirmation that the operation should continue .
 n phoning Weaver to announce his immediate availability , and **asking** for 'news'.19 Although Marton fumbled MIT by coming across as
 tive data such as credit card numbers , and authentication by **asking** for a password .
 their system administrators to use the address of the client **asking** for a page to control access to certain Web pages--such as th
 They could **ask** for soft drinks and meals (breakfast , lunch or dinner).
 r when the user closes the window or when the user explicitly **asks** for an " update . "

逆に (-1,1) の順で指定すると、検索語の左でソートして揃った行を、さらに検索語の右隣でソートすることになります。下図は直前の"also"で揃えて、次いで直後で揃えた結果になっています。

ソート設定 : 詳細ソート ▾

ソート詳細 : 優先度高: -1 ▾ 1 ▾ ▾ ▾ ▾ : 低

They were also **asked** about any other drugs they used that were not covered in thos
 Zimbabwe survey participatns were also **asked** questions regarding manure management , livestock , well cons
 Each patient was also **asked** their reason for stopping exercise immediately after each exe
 Subjects were also **asked** to provide a 30ml blood sample , which was sent by mail to th
 Owners of cases were also **asked** to describe any previous history of LUTD .
 aired performance in Experiments 1a and 1b , when he was also **asked** to judge the position of a part (a dot) relative_to a whole

※(3)左ソート,(4)右ソート,(5)中心語を含む右ソート、中心語を含む左ソートはソート詳細では以下のように設定して再検索した場合と同じになります。

	第1優先キー	第2優先	第3優先	第4優先	第5優先
(3) 左ソート	-1	-2	-3	-4	-5
(4) 右ソート	1	2	3	4	5
(5) 中心語を含む左ソート	0	-1	-2	-3	-4
中心語を含む右ソート	0	1	2	3	4

検索語と検索結果の位置を示す詳細検索の数値の対応関係

検索語(フレーズ)の例 make * up with

検索結果の各単語位置	-4	-3	-2	-1	0	1	2	3	4
検索結果(例)	you	own	a	machine	make	it	up	with	help
	,	they	should	sensibly	make	up	with	their	ex-enemy

(7) 「ソートキー」の設定

ソートキー:	<input checked="" type="checkbox"/> 表記形	<input type="checkbox"/> 品詞	<input type="checkbox"/> 基本形
	<input type="checkbox"/> 大・小文字を区別		

初期設定では、ソートの対象となる文字列は、実際に文に現われる語（表記形）になっています。ヨーパスのデータは、実際にはサーバー上では、表記形以外に品詞コードと基本形の3つ組みで取り扱っています。このことを利用して、ソートの対象となる文字列を変更すると、さらに詳細なソートを行うことができます。

たとえば、"agree to"で検索した結果には、to 不定詞と to が前置詞の例が混在しています。ソートキーに「品詞」を加えると、"to|T00"（不定詞用法）と"to|PRP"（前置詞）が文字列として区別されます。+1の位置でソートすると、まず"to|PRP"が揃い、その後に"to|T00"が揃います（品詞コードのアルファベット順に並びます）。切れ目の確認を簡単にするために、さらに"to"の右を2番目のキーにしておきます。次の図は、そのような仕様でソートキーの設定とソートの優先順位の設定を行い、検索した結果です。"agree to Zhu's elevation"で前置詞の用例が終了し、次の行"take in this"から不定法の用例が始まっていることがわかります。

なお、連語や*を使用した検索では、検索語の最初の単語がオレンジ文字、それ以外はボールドになりますが、キーの順番は、検索語(句)の最初の単語が0で、右隣が+1になり、あくまでも単語の切れ目で数えます。

ソート設定:	詳細ソート▼		
ソート詳細: 優先度高:	+1 ▼ +2 ▼ ▼ ▼ ▼ :低		
ソートキー:	<input checked="" type="checkbox"/> 表記形	<input checked="" type="checkbox"/> 品詞	<input type="checkbox"/> 基本形
	<input type="checkbox"/> 大・小文字を区別		

his discriminatory order was to be removed , the NWSs must agree to	a time-bound programme of nuclear disarmament .
There are also pressures on India to agree to	a Fissile Materials Cut-off Treaty (FMCT) and before tha
It is also not in India 's interest to agree to	a moratorium on the production of fissile materials unless
one adopted by all parties in 1874 , so that traders could agree to	a contract before seeing the commodity .
Analytical and numerical solutions agree to	a great extent in the near field region , leading to a mut
Each participant agreed to	an annual detailed clinical evaluation and brain donation
Fortunately , she agreed to	another postpartum study after resolution of her gastroint

(8) 「大文字/小文字を区別」

大文字と小文字を区別してソートします。検索は大文字と小文字を区別しませんが、この場合は、ASCII のコード順にソートしますので、記号類、数字、大文字、小文字の順に並びます。従って、検索後に小文字と大文字を KWIC 画面上で分けることができます。

5 KWIC 画面のダウンロード



KWIC 画面上部にある「ダウンロード」をクリックすると、画面に表示されたデータを CSV 形式でダウンロードすることができます。中心語の左右の文脈は、個別のフィールドが割当てられます。ただし「表示変更」で「1 文表示」が選ばれている場合には、中心語で分割されずに 1 文 1 行の形式となります。もし、CSV のファイルタイプが MS Excel と関連付けられている場合には、MS Excel を起動することもできます。

6 「クラスター」検索

A screenshot of the KWIC search results page with a detailed view of the 'Cluster' analysis settings. The main interface shows search results for 'Prepacking cooling ha...' and other entries. Overlaid on this is a 'Cluster' configuration dialog box. The dialog includes fields for '集計範囲' (Count Range) set to 0 ~ 0, '頻度' (Frequency) set to 2, and a 'クラスター' button. Below these are sections for '集計キー' (Count Key) with '表記形' checked, 'ソート' (Sort) with '頻度順' selected, and a '条件' (Conditions) section with 'この条件で語(句)集計を表示する' (Show Count by Word/Phrase under these conditions). The background shows the KWIC search results table.

「クラスター」は、検索後の KWIC の全行にわたり、行の特定の範囲を調べて、そこに同じ形で登場する語句を集計する機能です。この機能を使いこなすと、頻度の高い語句や表現を簡単にリストアップすることができます。「この条件で語（句）集計を表示する」を押すと設定画面が閉じてから処理が始まり、集計結果は別の画面に表示されます。

＜注意＞ 語句の集計は、BNC 全体ではなくて、検索結果の KWIC の全行に対する処理です。出現数の多い語を検索する場合、精度を確保するには、あらかじめ「表示設定」の「検索結果上限」を 10,000 件程度（もしくはそれ以上）にしてください。

■ 「集計範囲」

KWIC の各行のどの部分の語句を集計するかを指定します。集計の範囲は、始点と終点で指定します。指定の方法は、中心語を 0 の位置として、そこから数えて前方（画面の左方向）は、「-」符号付きの数字で、後方（画面の右方向）は「+」符号付きの数字で、語の位置を表します。例えば、中心語を挟んで、前後の 2 語を含む計 3 語で集計する場合は、「-1～+1」となります。また、中心語の後方 3 語を集計する場合は、「+1～+3」となります。

特定の場所の一語だけを集計したい場合は、「-2～-2」のように始点と終点を同じにします。一般的に、集計の範囲の長さを長くすると、異なり数が増えて、1 件あたりの頻度が減り、特徴のある集計結果とはなりませんから、段階的に長くしてください。また、同じ長さでも始点と終点を変えると、結果は当然異なります。

"make up" が取る目的語の頻度リストを得る課題をどのように解決するかを、以下に示します。

まず、目的語を 2 語のフレーズで調べたい場合には、図のように「集計範囲」を +2～+3 に設定します。+2～+3 の設定は、"make(0) up(+1) A(+2) B(+3)"（括弧内は語の位置）を意味し、"A B" の位置に出現する連語を集計することになります。



「この条件で語（句）集計を表示する」ボタンを押すと、この画面が消えて集計が始まり、次の図のように結果が表示されます。語句の長さがすべて 2 語になっています。

クラスター検索結果					
ダウンロード 語(句)検索結果ページに戻る					
検索	表記形	基本形	単語間の語数	出現位置	指定集計範囲
語(句)検索	-	make up	1 語以内	なし	+2 ~ +3 OFF
サブコードバスクール ALL					
#	頻度%	語句			
1	18	5.31 of a			
2	14	4.13 of the			
3	8	2.36 of two			
4	5	1.47 of several			
5	4	1.18 of local			
6	3	0.88 of many			
7	3	0.88 of n			
8	3	0.88 the bulk			
9	2	0.59 90% of			
10	2	0.59 a four-digit			
11	2	0.59 a large			
12	2	0.59 a substantial			
13	2	0.59 as a			
14	2	0.59 by the			
15	2	0.59 for by			

■ 「1単語で集計する」

指定した範囲について、その長さで集計するのではなく、単にその範囲に含まれる単語1語で集計したい場合には、下段の「1単語で集計」をチェックします。さらに、その中から特定の品詞を除外したい場合、あるいは、特定の品詞だけを残したい場合には、品詞コードで指定することができます。品詞は品詞一覧から複数を指定できます。

例えば、次の図は、"help"の検索結果から、+1～+3の範囲に登場する名詞を集計して、頻度順に基本形で表示させる例です。

集計範囲： +1 ~ +3 頻度： 2 以上を表示

集計キー： 表記形 品詞 基本形

ソート： 頻度順 アルファベット順

1単語で集計 品詞 N.* 含む 含まない

この条件で語(句)集計を表示する



クラスター検索結果

[ダウンロード](#) | [語\(句\)検索結果ページに戻る](#)

検索	表記形	基本形	単語間の語数	出現位置	指定	集計範囲	1単語で集計	集計品詞
語(句)検索	—	help	1語以内	なし	+1 ~ +3	ON	ON	N.* を含む
サブコーパス	ALL							

#	頻度%	語句
1	16	0.42 user
2	10	0.26 patient
3	10	0.26 people
4	9	0.24 designer
5	9	0.24 problem
6	8	0.21 system
7	7	0.18 developer
8	7	0.18 therapy
9	6	0.16 equation
10	6	0.16 physician
11	6	0.16 reader
12	6	0.16 requirement
13	6	0.16 scientist
14	6	0.16 understanding
15	5	0.13 project
16	5	0.13 research
17	4	0.10 behavior

■ 「品 詞」

品詞を指定するには、青字の「品詞」を押して品詞の一覧リストを出します。指定したい品詞を選んで「検索条件に追加する」クリックすると、その品詞コードがセットされます。品詞はOR条件でいくつでも同時指定することができます。

(1) 品詞一覧表の見方と利用方法

品詞分類は上下関係を持ちます。一覧表では、一番左が親で右に向って子供に細分されます。この性質を利用すると、最下位分類だけでなく、上位や途中の分類を利用して一度に複数の品詞を指定することができます。例えば、「名詞」を選ぶと、一度に「普通名詞」と「固有名詞」を含んだ指定となりますし、「普通名詞」を選ぶと、「固有名詞」は除外した単数、複数の名詞を全て指定したことになります。

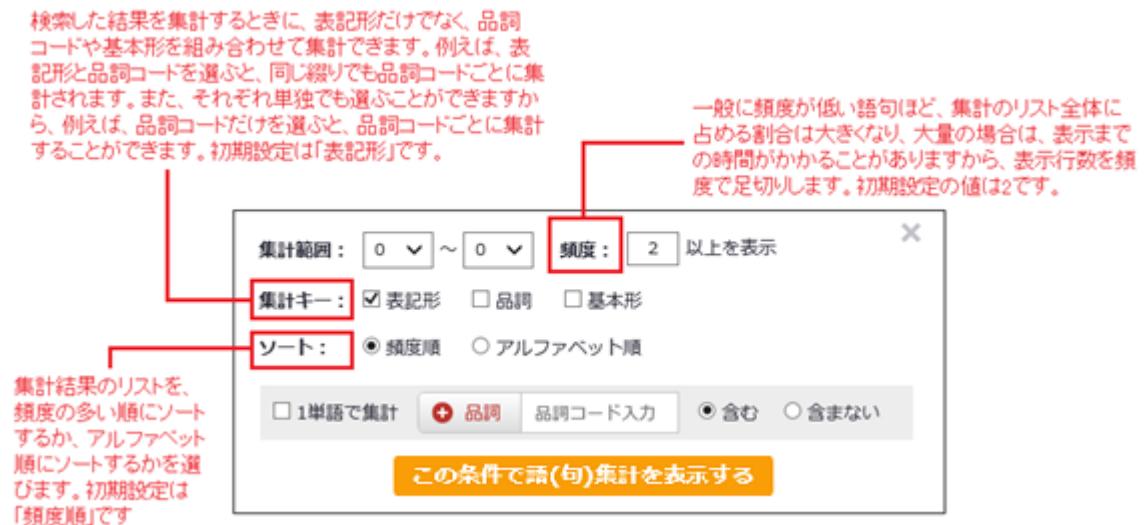
品詞一覧 品詞をお選びください。 検索条件に追加する

品詞				例
□ 名詞	□ 普通名詞	□ 単複両形	NN0	aircraft, data, committeeなど。単複両方扱いをされる単数形の名詞。
		□ 單数形	NN1	pencil, goose, time, revelationなど。
		□ 複数形	NN2	pencils, geese, times, revelationsなど。
	□ 固有名詞	NN0	London, Michael, Mars, IBMなど。単複の区別はない。	
□ 代名詞	□ 不定代名詞	PNI	none, everything, one, nobodyなど。someやtheseなどは含まれず、冠詞、決定詞（一般）に含まれることに注意。	
	□ 人称代名詞、所有代名詞	PNP	I, you, them, oursなど。yours, ours, hersなどの所有代名詞もここに含まれる。	
	□ 再帰代名詞	PNX	myself, yourself, itself, ourselvesなど。	
	□ 指示代名詞	DT0	thisなど。指示形容詞もここに含まれる。	
□ 形容詞	□ 形容詞一般	□ 原級	AJ0	good, oldなど
		□ 比較級	AJC	better, olderなど
		□ 最上級	AJS	best, oldestなど
	□ 所有形容詞	DPS	your, their, hisなど。	

(2) 異なる品詞を複数指定する方法

すでに「名詞」を指定した後で、さらに「形容詞一般」を指定する場合は、単純に順次、希望の品詞をクリックすれば、OR 条件として重ねることができます。OR 条件は、半角の垂直線“|”が自動的に挿入されます。

■ 「頻度」、「集計キー」、「ソート」



■ 「ダウンロード」

クラスター検索結果						
ダウンロード 語(句)検索結果ページに戻る						
検索	表記形	基本形	単語間の語数	出現位置指定	集計範囲	1単語で集計
語(句)検索	—	help	1語以内	なし	+1 ~ +3 ON	N.* を含む
サブコーパス	ALL					

#	頻度%	語句
1	16	help

クラスター検索の集計結果は、集計条件をヘッダーにした CSV 形式でダウンロードすることができます。

7 「分布グラフ」

「語(句)検索」の結果を、KWIC 表示だけでなく、グラフ表示で視覚的に表すことができます。



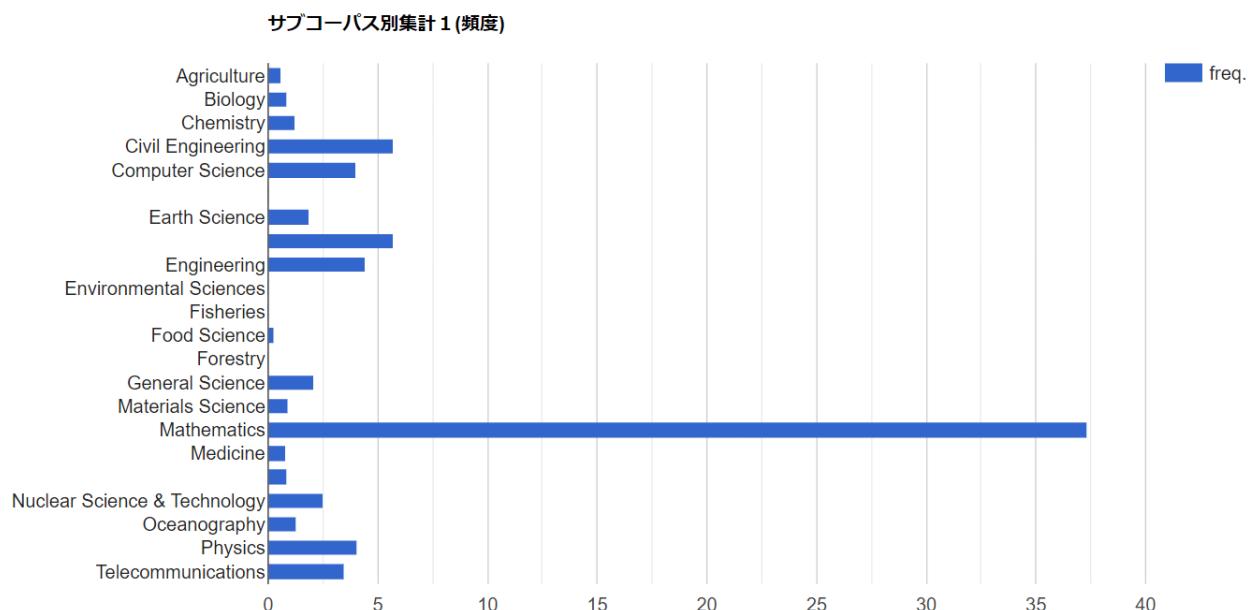
The screenshot shows the KWIC search interface with the following details:

- Top navigation bar: 語(句)検索, コロケーション, 單語情報, 初期設定, マニュアル.
- Search bar: converge, 検索 (Search) button.
- Search settings: 基本形で検索 (selected), サブコーパス [設定], 2単語前の語数: 指定なし, 2語以内.
- Section title: 語(句)検索結果.
- Search results: 475件 (2.93 / 10万語).
 - Row 1: ... which form the beginning of a subsequence which will eventually converge to &formula; from above .
 - Row 2: ... formula; ; if convergent to a nonhyperbolic fixed point of g , converges very slowly .
 - Row 3: ... ciated to g into an iterative process associated to &formula; ; converging faster .
 - Row 4: ... mptions , the iterative processes associated to NH and NC will converge faster to a fixed point (if one starts sufficiently close to
 - Row 5: ... Hereafter when we say , for instance , that &formula; ; converges to &formula; ; we mean that the associated iterative process &
 - Row 6: ... , we mean that the associated iterative process &formula; ; will converge to &formula; .
- Buttons: 分布グラフ ON (highlighted with a red oval), ソート・表示変更, クラスター, ダウンロード.
- Pagination: 1 / 19.

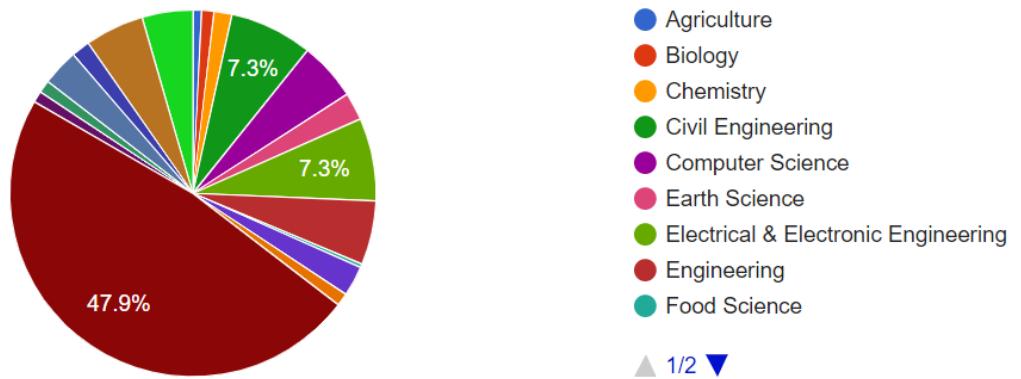
KWIC 画面上部にある「分布グラフ」は、初期設定で OFF になっています。ON の部分をクリックすると、ON が反転表示され、



KWIC 画面の下にグラフが表示されます。次の図は、PERC Corpus 全体で receipt を基本形で検索した後、「分布グラフ」を ON にして表示したものです。



サブコーパス別集計2(割合)



10万件の1あたりの出現件数に換算したサブコーパスごとの頻度が、棒グラフと円グラフで表示されます。棒グラフは出現頻度で、円グラフは出現率で表示しております。(棒グラフのサブコーパス名は、数が多くて表示しきれないときは、間引いて表示されていることがあります)

棒グラフは棒の部分に、円グラフは扇形の部分にカーソルを当てると、サブコーパス名と頻度、または出現%が表示されます。

8 「コロケーション」



「コロケーション」検索をしたい場合は、検索画面の「語(句)検索」、「コロケーション」、「単語情報」の3つのタブのうちから「コロケーション」をクリックします。

「コロケーション」検索は、ある語が文中でどの語と共に使用されているかについて、その語の前後5語の範囲にわたって頻度などを調べた表を表示します。表の中心に調べたい語が位置するので、**中心語**と呼び、前後に現われる語を**共起語**と呼びます。

■ 中心語の指定

入力エリアに調べたい共起関係の中心となる語を1語だけ入力します。中心語は、検索結果となる共起表の中心にレイアウトされます。「検索」ボタンを押すと検索が開始されます。使用できる文字種は、アルファベットと数字だけで、大文字と小文字は区別されません。記号は使用できません。

〈注意〉 「共起」の検索語に、連語は指定できませんから、複合語や句動詞など2語以上で1語として扱いたい語の共起を調べることはできません。

(1) 「基本形で検索」

語形変化をすべて含んで検索する場合に選びます。この場合、共起する語も全て基本形で集計される点にご注意ください。

(2) 中心語、共起語の品詞設定

両方の「品詞コード入力」エリアの左側にある、「中心語の品詞」、「共起語の品詞」をクリックすると、「品詞一覧」が表示されます。指定したい品詞を選んでクリックすると、その品詞コードが自動的に「品詞コード入力」エリアに入力されます。詳細は「7 「クラスター」検索」の「品詞」を参照してください。

■ 共起語の設定

共起語の表示は、設定を何も行わない場合、全ての品詞にわたって、共起頻度の多い順にソートしたリストとなります。中心語にどのような語を指定しても、大抵の場合、絶対的に頻度が高い冠詞や代名詞などが共起表の上位を常時占める結果となります。特定の品詞について調べたい場合に、これでは非常に不自由ですから、表示させたい品詞を指定できます。

また、共起語のソート方法は、単純な共起頻度以外に、3種類の統計量を用意しており、適宜、切り替えることができます。

〈注意〉 共起頻度とは、語 A と語 B が同じ文中で近傍に出現した回数です。初期設定は前後 3 語の範囲を個別に集計しています。詳しくは「共起表の見方」を参照してください。

■ 「共起語のソート」

共起語のリストをどの統計量にしたがってソートするかを選びます。初期設定は、共起頻度の絶対値の大きい順位です。



(1) 「T スコア」

T スコアは、統計の手法として知られる T 検定を応用したものです。T スコアでは帰無仮説として、ランダムな語の並びにおける共起確率が使用されています。したがって、T スコアが十分に高い場合は、偶然の共起ではないという判断が可能です。

(2) 「LogLog スコア」

M. I. の結果に、共起頻度の対数を乗じた値です。実際の共起頻度が値に反映されるので、後述の M. I. の欠点が解消されています。一般的には、名詞句の慣用的な表現を調べる場合などに、重宝します。

(3) 「M.I.スコア」

「M. I.」は、Mutual Information の略表記で、「相互情報量」と訳されます。M. I. は、共起頻度の確率を、2語の単独頻度の確率の積で割って対数化したものです。この値から2語がどれだけ独占的に共起しているかの程度を知ることができます。

単独の頻度に対して共起頻度が高ければ、M. I. の値は高くなります。ただし、共起頻度が極端に小さい場合で、かつ単独頻度も十分に小さい場合に、値が大きくなり過ぎることがありますから、上位にランクされた共起語の頻度には、必ず注意してください。

■ 「共起表」の見方

The screenshot shows a search interface for 'classical' with various filters and sorting options. The results table displays 100 rows of co-occurrence data, with columns for rank, words at -3..-1, -3, -2, -1, 0, 1, 2, 3, and 1..3 positions relative to the center word.

行数:100【表記形】 ソート:共起頻度									
Rank	-3 .. -1	-3	-2	-1	0	1	2	3	1 .. 3
1	the	488	the	42	of	72	the	408	classical 2018
2	of	146	,	39	in	56	a	85	and 23
3	in	118	of	24	,	44	in	51	theory 23
4	a	106	to	17	to	41	of	50	*
5	,	93	we	14	the	38	and	21	mechanics 15
6	to	73	a	13	with	25	with	21	gompertz 14
7	with	46	be	13	is	19	from	19	linear 13
8	and	35)	12	for	17	*	16	model 10
9	by	35	is	12	on	17	by	15	moduli 10
10	is	34	in	11	as	15	to	15	approach 9

共起表の列は、検索語を中心とした前後3語の範囲と、さらにその外側に前後の範囲のデータをマージした計9列から構成されます。

行は、その位置での共起語のリストです。行は、統計量の多い順にソートされます。統計量は、共起頻度以外に、Tスコア、M.I.スコア、LogLogスコアの計4種類があります。

共起表は、各列ごとに集計した結果ですから、列と列の間にどのような統語的な関係があるかをることはできません。つまり、隣り合った列の、リストの1位に来ている2語が、連語としてやはり1位となるかどうかについて、共起表から知ることはできません。あくまでも、中心語と各列の2語の共起関係を調べた結果です。

(1) 出現頻度

中心語の下に表示されるのは、その語の PERC 中の総頻度数です。

-1	0	1
the 408	classical 2018	and 23
a 85		theory 23
in 51		" 15
of 50		mechanics 15
and 21		gomPERTZ 14
with 21		linear 13
from 19		model 10

共起語の統計値は、共起語の右に表示されます。共起頻度でソートした場合は、共起頻度だけが表示されます。他の統計量でソートした場合には、その値に続けて（ ）内に共起頻度も表示されます。特に M. I. でソートしたときには、上位の共起頻度に注意してください。

下図は、M. I. でソートした例です。なお、各列とも、品詞を無視した場合の出現頻度が 1 以下の単語 g は集計していません。

cell のコロケーション ダウンロード

◀ 1 / 4 ▶

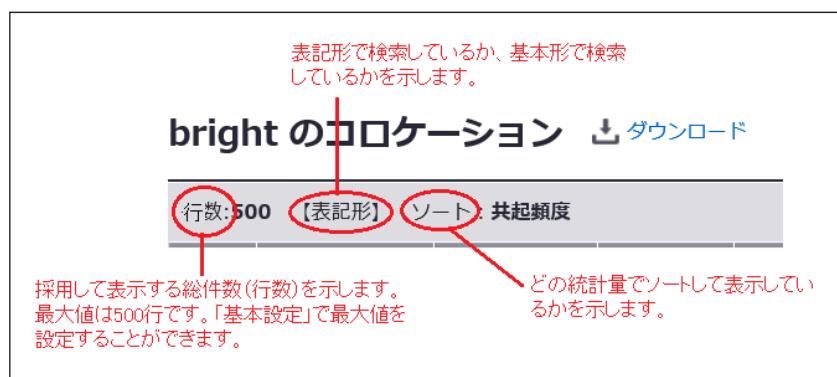
行数:100【表記形】 ソート: M.I. スコア ※ ()内共起頻度																
Rank	-3 .. -1		-3		-2		-1		0	1	2	3				
1	carcinoma-derived	9.74 (4)	follicular/h	9.74 (4)	formed/min/mg	9.74 (3)	carcinoma-derived	9.74 (4)	cell 24134	line-derived	9.74 (7)	besc	9.74 (2)	extrathymic	9.32 (3)	t
2	cell-golgi	9.74 (3)	prestalk	7.01 (8)	chv-infected	9.32 (3)	cell-golgi	9.74 (3)		scraper	9.74 (4)	diachema	9.74 (2)	glutathione-px	8.15 (3)	dia
3	cell-t	9.74 (3)	numa	6.44 (6)	lio[91-99]-specific	9.15 (4)	cell-t	9.74 (3)		surface-associated	9.74 (11)	.0.5	8.93 (4)	20.4	7.23 (3)	line-
4	element-companion	9.74 (3)	commit	6.37 (3)	a431	8.32 (3)	element-companion	9.74 (3)		corpse	9.54 (7)	liferation	8.93 (4)	huly	7.15 (3)	sc
5	fluorescence-activated	9.74 (5)	detached	6.00 (3)	basophilic	8.15 (11)	fluorescence-activated	9.74 (5)		sorter	9.32 (3)	coexpressing	8.74 (2)	v[rbc]	7.15 (4)	su assi
6	follicular/h	9.74 (4)	diffuses	6.00 (3)	encephalitogenic	8.15 (3)	impaled	9.74 (4)		harvester	9.20 (9)	tiation	8.32 (3)	fac	6.98 (4)	cr
7	formed/min/mg	9.74 (3)	aiming	5.86 (3)	undivided	8.00 (3)	leukocyte-endothelial	9.74 (3)		patch-clamp	9.15 (4)	extensibility	8.15 (3)	cytokine	6.41 (6)	extr
8	impaired	9.74 (4)	cdk5	5.62 (3)	cd8+	7.43 (17)	non-small	9.74 (5)		hybridomas	9.11 (11)	agnor	7.93 (2)	epiphytic	6.28 (3)	s
9	leukocyte-endothelial	9.74 (3)	aggregated	4.98 (3)	mesodermal	6.93 (3)	rthle	9.74 (5)		differen-	8.74 (3)	depolarisation	7.93 (2)	axonal	6.22 (7)	har
10	non-small	9.74 (5)	transported	4.91 (10)	antigen-specific	6.84 (7)	sickle	9.64 (15)		prolifer-	8.74 (3)	asynchrony	7.74 (2)	int	6.18 (4)	patc

(2) 「-3～-1」と「+1～+3」の見方と集計方法

検索語に対して、知りたい共起語の位置が不定の場合があります。例えば、"want"の目的語に相当する語としてどのような名詞が共起しているかを概観したい場合には、「+1～+3」の列を見ます。ただし、すべての例が、want の目的語とは限りませんので、実際の例文で必ず確認してください。

「-1～-3」および「+1～+3」の値は、それぞれ、-1、-2、-3（または+1、+2、+3）の各列に出現した同じ語について、その共起頻度を合算した共起頻度、およびそれに基づいて他の統計量を計算しています。

(3) 共起表最上段の数値表示について



■ 表示設定

検索画面上部にある「初期設定」をクリックして、共起表の表示設定を行うことができます。



設定できるのは、「検索結果上限」（初期設定は 100 行）と「1 ページ行数」（初期設定は 25 行）です。入力したあとに「設定を保存する」ボタンを押してください。



■ KWIC 画面表示

共起表の共起語をクリックすると、実際の用例を KWIC 画面に表示することができます。このとき、「初期設定」の「語（句）検索」の中にある「検索結果上限」で、サーバーから検索する上限数を設定することができます。初期設定は 3000 行です。なお、最大は 2 万行です。KWIC 画面の見方や表示の設定方法などは、「[6 KWIC 画面のダウンロード](#)」をご覧ください。

■ 「ダウンロード」

共起表を CSV 形式でダウンロードします。CSV のファイルタイプが MS Excel に関連付けられている場合には、MS Excel で開くこともできます。「[6 KWIC 画面のダウンロード](#)」をご覧ください。

9 「単語情報」

「単語情報」は、検索語を入れると、BNC 中での品詞別の頻度表を表示します。調べたい語の品詞が不明の場合にも利用できます。

次の図は、“forget”を基本形で検索した例です。検索語の品詞を指定したい場合は、「品詞指定」をクリックして、「品詞一覧」の表から指定を行うことができます。ここでは品詞は 1 種類しか指定できません。

語(句)検索 コロケーション 単語情報

初期設定 マニュアル

品詞指定 品詞コード入力 指定した品詞を 含む

forget 基本形で検索 検索条件 完全一致 表示件数/ページ 100

件数:1件 【基本形】

No.	基本形	基本形合計	表記形	表記形合計	品詞別頻度
1	forget	56	forgotten	28	VVN 28
			forgetting	14	VVG 14
			forget	11	VVI 8
			forgot	3	VVD 3

■ 検索条件

検索したい語は、「完全一致」「前方一致」「後方一致」「部分一致」の4つの検索方法で検索することができます。

語(句)検索 コロケーション 単語情報

品詞指定 品詞コード入力 指定した品詞を 含む

単語または単語の一部を入力して下さい

基本形で検索 検索条件 表示件数/ページ 100

検索条件 フラグ

- 完全一致
- 完全一致
- 前方一致
- 後方一致
- 部分一致

cell の単数形(NN1)を後方一致で検索した例

語(句)検索 コロケーション 単語情報

品詞指定 NN1 指定した品詞を 含む

cell 基本形で検索 検索条件 後方一致 表示件数/ページ 100

件数:70件 【表記形】

No.	基本形	基本形合計	表記形	表記形合計	小計	品詞別頻度
1	cell	29863	cells	17796	0	NN2 17796
			cell	12067	12067	NN1 12067
2	t-cell	573	t-cell	426	426	NN1 426
			t-cells	147	0	NN2 147
3	microcell	146	microcell	97	93	NN1 93
			microcells	49	0	NN2 45
4	b-cell	119	b-cell	98	12	AJ0 86
			b-cells	21	0	NN2 21
5	whole-cell	96	whole-cell	96	6	AJ0 89
						NN1 6
						NPO 1
						VVI 1

■ KWIC 表示と原文表示

検索結果の一覧のうち、数字や品詞が色文字になっています。これらをクリックすると KWIC が表示されます。KWIC の中心語をクリックすると原文が表示されます。また、KWIC 情報をダウンロードすることもできます。これらの機能や表示設定などについても、「語(句)検索」、「コロケーション」における KWIC の機能、表示設定と同様ですので、参照してください。

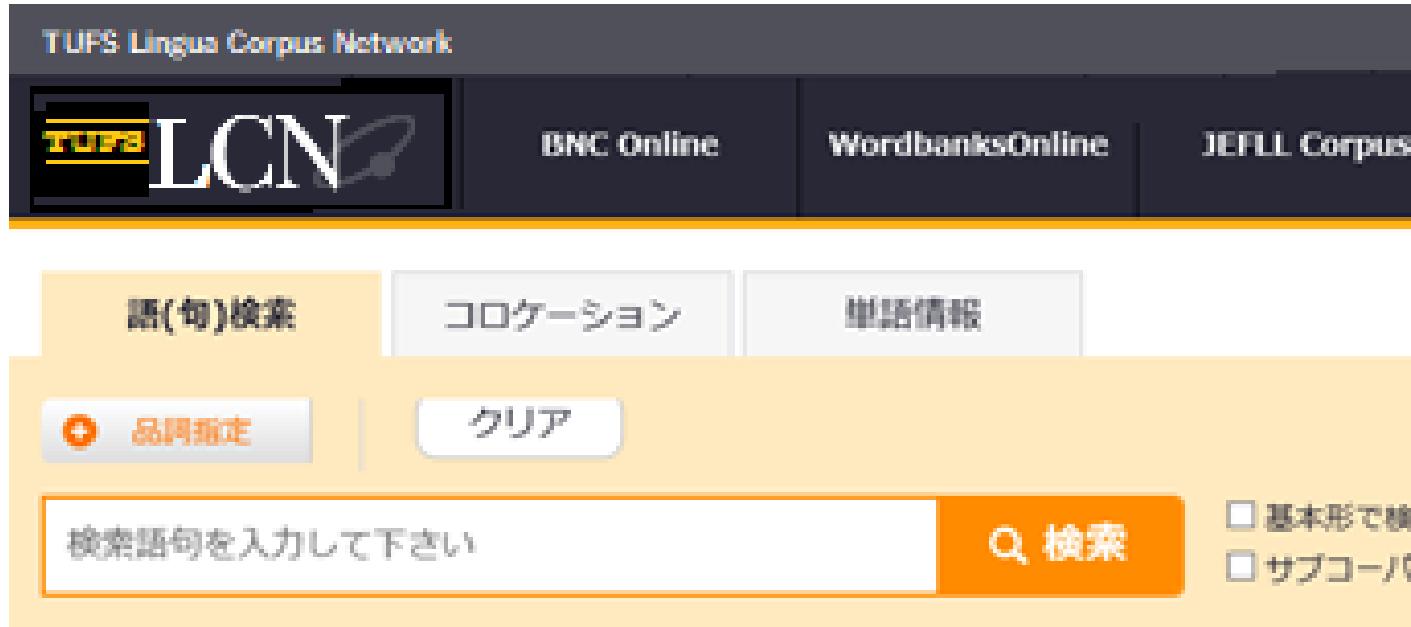
10 「ログアウト」

ブラウザーからの検索を一旦終了される場合は、必ず以下のやり方で「ログアウト」を行ってください。

■ ログアウトの方法

正しいログアウトの方法を説明します。

使用後は**ブラウザーの「ファイル」メニューにある「閉じる」やウィンドウ外枠の右上隅の「×**で**ブラウザーを終了させないでください**。図に示すように、検索画面の右上にある「ログアウト」ボタンを押して終了してください。



■ 「ログアウト」ボタンを押さずに、ブラウザーを閉じた場合に起こること

ブラウザーの「×などで終了した場合は、サーバー側ではまだログインしたままの状態であると判断しています。しかし、そのまま 30 分が経過しますと、サーバー側で自動的にログアウトします。これを「タイムアウト」と呼びます。

もしタイムアウトの前に別のマシンや同じマシンからログインすると、「利用できるアクセス数を超えています」とアクセス数がご契約のID数を超えている警告が出てログインできません。これは、不正使用を避ける機構が正常に働いているためです。



■ なんらかの障害で正常にログアウトできなかった場合に起こること

回線が切れたり、マシンやブラウザーが異常終了した場合も、上記と同じで正常にログアウトしていないため、タイムアウトの前には、別のマシンからはログインできません。異常終了によりクッキーのデータが壊れてしまった場合には、タイムアウトの前でも同じマシンからログインできないことがあります。この場合も、サーバー側でログアウトするまでの約30分程度お待ちください。

TUFS Lingua Corpus Network
PERC Corpus 検索ソフト使用マニュアル 2024年度9月版

東京都府中市朝日町3-11-1
東京外国語大学ワールド・ランゲージ・センター
©TUFS World Language Centre 2024

本書の一部あるいは全部を無断で複製・転載することは法律で認められた場合を除き、著作者の権利の侵害となります。あらかじめ当センターにて許諾を求めてください。