

テーマ 1	複数のコーパスに共通してよく使用される語																																																																		
	<p>ほぼすべてのコーパスに出現するけれども、2つまたは3つのコーパスに突出して多くみられる語を探してください。また、それがなぜなのか、それらのコーパスに共通していることは何なのかなど、考えられる理由を書いてください。</p> <p>(検索対象：コーパス毎)</p>																																																																		
提出項目	(複数回答可) [必須] 検索条件 (検索フォームへの入力内容とスクリーンショット) [必須] 検索結果 (棒グラフ) [必須] 考察																																																																		
提出項目例	<ul style="list-style-type: none"> 検索条件 書字形出現形「よろしく」+書字形出現形「お」+書字形出現形「願い」+語彙素「為る」+語彙素「ます」 <div data-bbox="459 967 1102 1541" data-label="Form"> </div> <ul style="list-style-type: none"> 検索結果 <div data-bbox="459 1603 1369 1966" data-label="Figure"> <table border="1"> <thead> <tr> <th>カテゴリ</th> <th>コーパス</th> <th>説明</th> <th>検索結果の件数</th> <th>検索対象語数</th> <th>状態</th> </tr> </thead> <tbody> <tr> <td>BCCWJ</td> <td>BCCWJ</td> <td>全てのデータが対象</td> <td>2,698</td> <td>104,911,460</td> <td>成功 検索結果を見る</td> </tr> <tr> <td>NWJC</td> <td>NWJC</td> <td>一部のデータが対象</td> <td>2,967</td> <td>86,277,772</td> <td>成功 検索結果を見る</td> </tr> <tr> <td>CSJ</td> <td>CSJ</td> <td>全てのデータが対象</td> <td>174</td> <td>7,576,046</td> <td>成功 検索結果を見る</td> </tr> <tr> <td>CEJC</td> <td>CEJC</td> <td>全てのデータが対象</td> <td>25</td> <td>1,247,080</td> <td>成功 検索結果を見る</td> </tr> <tr> <td>SSC</td> <td>SSC</td> <td>全てのデータが対象</td> <td>5</td> <td>528,589</td> <td>成功 検索結果を見る</td> </tr> <tr> <td>NUCC</td> <td>NUCC</td> <td>全てのデータが対象</td> <td>18</td> <td>1,131,971</td> <td>成功 検索結果を見る</td> </tr> <tr> <td>CWPC</td> <td>CWPC</td> <td>全てのデータが対象</td> <td>21</td> <td>186,906</td> <td>成功 検索結果を見る</td> </tr> <tr> <td>CHJ</td> <td>CHJ</td> <td>全てのデータが対象</td> <td>0</td> <td>17,604,526</td> <td>成功</td> </tr> <tr> <td>COJADS</td> <td>COJADS</td> <td>全てのデータが対象</td> <td>7</td> <td>769,394</td> <td>成功 検索結果を見る</td> </tr> <tr> <td>I-JAS</td> <td>I-JAS</td> <td>全てのデータが対象 (日本語母語話者は除く)</td> <td>477</td> <td>3,667,903</td> <td>成功 検索結果を見る</td> </tr> </tbody> </table> </div> 	カテゴリ	コーパス	説明	検索結果の件数	検索対象語数	状態	BCCWJ	BCCWJ	全てのデータが対象	2,698	104,911,460	成功 検索結果を見る	NWJC	NWJC	一部のデータが対象	2,967	86,277,772	成功 検索結果を見る	CSJ	CSJ	全てのデータが対象	174	7,576,046	成功 検索結果を見る	CEJC	CEJC	全てのデータが対象	25	1,247,080	成功 検索結果を見る	SSC	SSC	全てのデータが対象	5	528,589	成功 検索結果を見る	NUCC	NUCC	全てのデータが対象	18	1,131,971	成功 検索結果を見る	CWPC	CWPC	全てのデータが対象	21	186,906	成功 検索結果を見る	CHJ	CHJ	全てのデータが対象	0	17,604,526	成功	COJADS	COJADS	全てのデータが対象	7	769,394	成功 検索結果を見る	I-JAS	I-JAS	全てのデータが対象 (日本語母語話者は除く)	477	3,667,903	成功 検索結果を見る
カテゴリ	コーパス	説明	検索結果の件数	検索対象語数	状態																																																														
BCCWJ	BCCWJ	全てのデータが対象	2,698	104,911,460	成功 検索結果を見る																																																														
NWJC	NWJC	一部のデータが対象	2,967	86,277,772	成功 検索結果を見る																																																														
CSJ	CSJ	全てのデータが対象	174	7,576,046	成功 検索結果を見る																																																														
CEJC	CEJC	全てのデータが対象	25	1,247,080	成功 検索結果を見る																																																														
SSC	SSC	全てのデータが対象	5	528,589	成功 検索結果を見る																																																														
NUCC	NUCC	全てのデータが対象	18	1,131,971	成功 検索結果を見る																																																														
CWPC	CWPC	全てのデータが対象	21	186,906	成功 検索結果を見る																																																														
CHJ	CHJ	全てのデータが対象	0	17,604,526	成功																																																														
COJADS	COJADS	全てのデータが対象	7	769,394	成功 検索結果を見る																																																														
I-JAS	I-JAS	全てのデータが対象 (日本語母語話者は除く)	477	3,667,903	成功 検索結果を見る																																																														

- 考察

「よろしくお願いします」という表現は、依頼場面などで使用される表現であるが、CWPC「現日研・職場談話コーパス」は職場の会話であるために依頼場面が多くみられるのではないか。また I-JAS については、ロールプレイで依頼場面を扱っており多く出現すると考えられる。

テーマ 2	一貫した歴史的増加・減少																																												
<p>「時代」別の検索をしたときの検索結果（調整頻度）の棒グラフの頂点が一貫して増加または減少している語や表現を見つけてください。また可能であればそのようになる理由について考察を加えてください。</p> <p>なお、以下のものは対象外とします。</p> <p>① 用例が出現しない（検索結果の件数がゼロの）時代があるもの</p> <p>② KOTOHOHA 検索コンテスト 2020 優秀賞受賞作品と同じ検索条件のもの</p> <p>(検索対象：時代)</p>																																													
提出項目	<p>(複数回答可)</p> <p>[必須] 検索条件 (検索フォームへの入力内容とスクリーンショット)</p> <p>[必須] 検索結果 (棒グラフ)</p> <p>[任意] 考察</p>																																												
提出項目例	<ul style="list-style-type: none"> ● 検索条件 <p>語彙素が「花」</p>  ● 検索結果  <table border="1" data-bbox="991 1451 1361 1809"> <thead> <tr> <th>カテゴリ</th> <th>コーパス</th> <th>説明</th> <th>検索結果の件数</th> <th>検索対象語数</th> </tr> </thead> <tbody> <tr> <td>奈良</td> <td>CHJ</td> <td>時代名「1奈良」のデータが対象</td> <td>321</td> <td>116,928</td> </tr> <tr> <td>平安</td> <td>CHJ</td> <td>時代名「2平安」のデータが対象</td> <td>1,833</td> <td>1,069,846</td> </tr> <tr> <td>鎌倉</td> <td>CHJ</td> <td>時代名「3鎌倉」のデータが対象</td> <td>496</td> <td>872,005</td> </tr> <tr> <td>室町</td> <td>CHJ</td> <td>時代名「4室町」のデータが対象</td> <td>164</td> <td>358,419</td> </tr> <tr> <td>江戸</td> <td>CHJ</td> <td>時代名「5江戸」のデータが対象</td> <td>338</td> <td>809,083</td> </tr> <tr> <td>明治・大正・昭和 (前期)</td> <td>CHJ</td> <td>時代名「6明治」「7大正」「8昭和」のデータが対象</td> <td>3,707</td> <td>13,640,028</td> </tr> <tr> <td rowspan="2">昭和 (後期)・平成</td> <td>BCOWJ</td> <td>全てのデータが対象</td> <td>17,872</td> <td>104,911,460</td> </tr> <tr> <td>SSC</td> <td>全てのデータが対象</td> <td>2</td> <td>180,280</td> </tr> </tbody> </table> ● 考察 <p>時代が下るにつれて「花」の調整頻度が低下するのは、… (略)</p> 	カテゴリ	コーパス	説明	検索結果の件数	検索対象語数	奈良	CHJ	時代名「1奈良」のデータが対象	321	116,928	平安	CHJ	時代名「2平安」のデータが対象	1,833	1,069,846	鎌倉	CHJ	時代名「3鎌倉」のデータが対象	496	872,005	室町	CHJ	時代名「4室町」のデータが対象	164	358,419	江戸	CHJ	時代名「5江戸」のデータが対象	338	809,083	明治・大正・昭和 (前期)	CHJ	時代名「6明治」「7大正」「8昭和」のデータが対象	3,707	13,640,028	昭和 (後期)・平成	BCOWJ	全てのデータが対象	17,872	104,911,460	SSC	全てのデータが対象	2	180,280
カテゴリ	コーパス	説明	検索結果の件数	検索対象語数																																									
奈良	CHJ	時代名「1奈良」のデータが対象	321	116,928																																									
平安	CHJ	時代名「2平安」のデータが対象	1,833	1,069,846																																									
鎌倉	CHJ	時代名「3鎌倉」のデータが対象	496	872,005																																									
室町	CHJ	時代名「4室町」のデータが対象	164	358,419																																									
江戸	CHJ	時代名「5江戸」のデータが対象	338	809,083																																									
明治・大正・昭和 (前期)	CHJ	時代名「6明治」「7大正」「8昭和」のデータが対象	3,707	13,640,028																																									
昭和 (後期)・平成	BCOWJ	全てのデータが対象	17,872	104,911,460																																									
	SSC	全てのデータが対象	2	180,280																																									

テーマ 3	一番長い複合名詞
<p>日本語では多くの名詞が連なってひとつの複合名詞を形成します。できるだけ長い複合名詞を見つけてください。また長い複合名詞が生じやすい条件ないし環境についての考察があれば、追記してください。</p> <p>複合語の長さの測り方にはいろいろな方法がありますが、今回は短単位数で数えることにします。「編集 会議」は2短単位、「全国 警察 本部長 会議」は5単位です（縦線は短単位の境界です）。</p> <p>複合名詞には、「ボランティア 井戸 端 会議」、「鶴見 まちづくり 推進 会議」、「OECD 厚生 大臣 会議」などのように、漢字以外に平仮名やカタカナやアルファベットを含んでも構いませんが、以下のものは対象外とします。</p> <p>① 「企業 と 防災 に 関する 検討 会議」、「持続 可能 な 開発 に 関する 世界 首脳 会議」などのように助詞や動詞や助詞相当表現（「及び」「並びに」など）を含むもの</p> <p>② 「アジア 太平 洋 経済 協力 会議（ APEC ） 貿易 相 会議」、「亀山 市 レジ 袋 削減（ 有料 化 ） ・ マイ バッグ 推進 会議」などのように補助記号を含むもの</p> <p>③ 「東京・埼玉・神奈川・千葉・群馬」「総務省・法務省・外務省・財務省・文部科学省・厚生労働省…」などのように複数の名詞を並列したもの</p> <p>（検索対象：コーパス毎）</p>	
提出項目	<p>（1件のみ）</p> <p>[必須] 検索条件（検索フォームへの入力内容とスクリーンショット）</p> <p>[必須] 検索結果（コーパス名、サンプルID、開始位置、連番、複合名詞のテキストとその短単位数）</p> <p>[任意] 考察</p>
提出項目例	<ul style="list-style-type: none"> 検索条件 <ul style="list-style-type: none"> キーを「文末・節単位末・発話単位末から1語」として、語彙素「会議」「組合」「係」「課」「部」「センター」「窓口」「担当」「茶」でそれぞれ検索し、個々の用例に含まれる短単位数を数えた。

KOTONOHA 検索コンテスト 2021 規定テーマ

前方共起条件を追加する

キー 1 キーの条件を指定しない


が

後方共起条件を追加する

検索対象

コーパス毎 書き言葉・話し言葉 時代

- 検索結果
上記の検索条件で最も長い複合語は以下の例であった。
コーパス名：BCCWJ
サンプル ID：OW6X_00079
開始位置：63700
連番：40350
複合名詞：第|3|回|アジア|防災|センター|専門|家|会議 (9 単位)
- 考察
長い複合名詞が生じやすいのは、… (略)

テーマ 4	各地で共通の語形が使われている動植物の名前
	<p>日本各地に広く生息しているにも関わらず、共通の語形が使われている動植物名を探してください。可能であれば、なぜそのような結果になるのか、気付いたことの説明も加えてください。</p> <p>動植物の名前は方言間でのバリエーションが多く、多様な語形が存在することが知られています（蜻蛉：トンボ、アキズ、アゲ、ケンゾ、ボイなど）。一方、日本各地に広く生息しているにも関わらず、共通の語形が使われている動植物も存在します。より多くの地点で使用されている語を見つけてください。提出項目例にあるように、「ウマ～ンマ～オマ」等の発音の違いは考慮せず、全て共通の語形とみなします。</p> <p>(検索対象：COJADS)</p>
提出項目	<p>(複数回答可)</p> <p>[必須] 検索条件 (検索フォームへの入力内容とスクリーンショット)</p> <p>[必須] 地点名・検索キーに対応する方言テキスト</p> <p>[任意] 考察</p>
提出項目例	<ul style="list-style-type: none"> ● 検索条件 語彙素「馬」 <div data-bbox="483 1173 1112 1332" style="border: 1px solid black; padding: 5px; margin: 10px 0;">  </div> <ul style="list-style-type: none"> ● 地点名・検索キーに対応する方言テキスト <ol style="list-style-type: none"> 1. 北海道中川郡豊頃町：ンマ 2. 青森県北津軽郡市浦村：マ 3. 青森県弘前市：ウマ、ンマ、マ 4. 岩手県遠野市：ウマ、オマ <p>(略)</p> <ol style="list-style-type: none"> 31. 熊本県球磨郡錦町：ウマ、ンマ 32. 大分県宇佐市：ウマ 33. 大分県大分郡挾間町：ウマ 34. 沖縄県国頭郡今帰仁村：ヌマ、マ