

ロシア語 ナショナルコーパス を使ってみよう！

【ロシア語ナショナルコーパス(英語版)】



◇ ロシア語ナショナルコーパスとは.....

ロシア語ナショナルコーパス(Национальный корпус русского языка, URL: <https://ruscorpora.ru/en/> (英語版))とは、ロシア科学アカデミーが作成した、テキストや発話が大規模に集積された無料で利用できるデータベースです。このコーパスは、2003年に公開されました。文学小説や新聞などのテキストからデータを収集し、現在(2025年4月)では、20億トークン(テキストを意味のある最小単位に分割したもの)以上が収録されています。収録されたテキストや発話には、形態的な情報や意味情報などが付与されています。

◇ ロシア語ナショナルコーパスができること.....

● ロシア語の例文がほしい。

とにかく様々なテキストがあります。「動詞が最初に来る例文が欲しい!」、「与格と対格が出てくる例文が欲しい!」...等、詳細な検索ができます。言語学のレポートや論文だけでなく、言語学習にも!

● ロシア語の方言差を知りたい。/ ロシア語と他の言語を比較したい。

/古いロシア語を知りたい。

方言コーパスや、同じ文に対して他の言語と比較ができるパラレルコーパス、歴史コーパスなど様々なコーパスがロシア語ナショナルコーパスで閲覧できます。メインコーパスは18世紀半ばから20世紀半ばまでのテキストが閲覧でき、また年代別に検索できます。

◇ コーパスの検索の仕方.....

ここではメインコーパス(Main corpus)の検索画面を掲載していますが、ロシア語ナショナルコーパス内の他のコーパスでも検索の仕方は大方同じです。

①Word: 語

コーパスは Word | の検索窓に一語ずつ入力して検索します。右端のプラス(+) ボタンを押すと、100 語以上増やせます。

②Lemma:

レンマ/レマ

この検索窓には、検索結果として出力される例文に含まれる語を入力します。Lemma では、入力した語の活用形、あるいは曲用形も含めて検索されます。

③Word : ワード

この検索窓にも、検索結果として出力される例文に含まれる語を入力します。Lemma と違い、Word はその入力語の語形で検索されます。

The screenshot shows the search interface for the Russian National Corpus. It features a search bar at the top, followed by options for 'Exact search', 'Collocation search', and 'Lemmas and tags search'. Below these are two columns for 'Word 1' and 'Word 2', each with fields for 'Lemma', 'Word', 'Gramm. features', 'Semantic features', 'Syntactic features of a word', and 'Addit. features'. A 'Distance' field is also present. At the bottom, there are 'Search' and 'Reset' buttons. The footer contains contact information and logos for the Russian National Corpus and Yandex.

④Gramm. features:

人称、性、数、格...といった情報を入力します。select をクリックし、チェックボックスにチェックを入れましょう。

⑤Semantic Features:

動物、道具、食べ物...といった情報を入力します。select をクリックし、チェックボックスにチェックを入れましょう。

⑥Syntactic features of a word:

主節、従属節にある...といった情報を入力します。select をクリックし、チェックボックスにチェックを入れましょう。

〈続き〉

⑦ Addit. features:

文頭・文末にある…といったそのほかのといった情報を入力します。select をクリックし、チェックボックスにチェックを入れましょう

⑧ Distance:

語と語の間の距離を入力します。例えば человек(人間)と хлеб(パン)を入力してみます。

Word 1 (X) Word 2 (X)

Lemma (🔍) (?) (X) Lemma (🔍) (?) (X)

человек хлеб

- Distance に from: 1 to: 1 を入力した場合…

2. Евгений Лукин. Клопики (2013) ↩

А русского человека хлебом не корми — дай подглядеть, чем сосед занимается. 📄 <…> 🗨

человек と хлеб が隣に並びます。

- Distance に from: 1 to: 2 を入力した場合…

1. митрополит Антоний (Блум). О жизни христианской (1990) ↩

Технически он это сделал: человек съел хлеб, съел суп, получил полтинник и ушел. 📄 <…> 🗨

человек と хлеб が съел(食べた)を挟んでいます。
このように 1 語ずつ距離を指定することができます。

⑨ Search:

①から⑧まで必要な情報を入力したら Search をクリック、またはタップしましょう。例文が画面上にずらっと並ぶ concordance や統計、頻度なども同時に検索することができます。

⑩ Reset:

今まで行った入力を全て削除します。

◇ 検索結果の見方.....

検索欄に入力ができたら、実際に検索してみましょう。こんな画面が出てきます。



今回は、concordance ページの例文について紹介します。

2. Евгений Лукин. Клопики (2013) ①

③ ④ ⑤

② А русского *человека хлебом* не корми — дай подглядеть, чем сосед занимается. ③ ④ ⑤

①例文番号/著者名/テキストタイトル/発表年:

ここでは、例文番号は2、著者名はЕвгений Лукин、テキストタイトルはКлопики、発行年は2013年と書いてあります。このあたりをクリック・タップすると、この例文の原典についての情報が閲覧できます。

②例文:

検索した条件に従った例文が表示されます。человек(人間)とхлеб(パン)を検索欄に入力したので、キーワードがオレンジで表示されています。キーワードをクリック・タップすると、その語の文法情報(格や数、性など)が確認できます。

③Copy Text:

例文、著者名、テキストタイトルをまとめてコピーすることができます。

④Show extended context:

この例文がある前後の文章を表示します。

⑤Show the syntactic structure

この例文の主述関係や修飾関係を図式化してくれます。

この冊子で紹介した以外の機能はたくさんあるので、ぜひ色々試してみてください。
わからないことがあれば、東京外国語大学附属図書館学習相談デスクへ！