『中納言』の概要

大村舞・石本祐一・岡照晃　（国立国語研究所　コーパス開発センター）

コーパス検索システム『中納言』

中納言はコーパスを単語や品詞などの**形態論情報**で検索可能な検索システム
（要ユーザ登録）

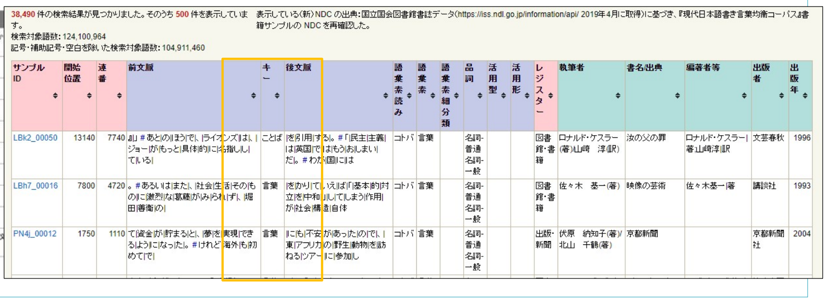
９つのコーパスについて検索可能

【個別検索】ご利用になりたいコーパス名をクリックしてください。

コーパス名	略称	個別検索	包括的検索	備考
書き言葉 現代日本語書き言葉均衡コーパス 中納言版	BCCWJ	✓	✓	従来より利用いただいている BCCWJ のデータです（コーパスの紹介ページ）。こちらのページから BCCWJ アノテーションデータをダウンロードできます。梵天版はこちら
書き言葉 国語研日本語ウェブコーパス 中納言版	NWJC	準備中	✓	梵天版はこちら
話し言葉 日本語話し言葉コーパス	CSJ	✓	✓	コーパスの紹介ページ
話し言葉 日本語日常会話コーパス モニター公開版	CEJC	✓	✓	コーパスの紹介ページ
話し言葉 昭和話し言葉コーパス モニター公開版	SSC	✓	✓	コーパスの紹介ページ
話し言葉 名大会話コーパス	NUCC	✓	✓	コーパスの紹介ページ
話し言葉 現日研・職場談話コーパス	CWPC	✓	✓	コーパスの紹介ページ
通時 日本語歴史コーパス	CHJ	✓	✓	コーパスの紹介ページ
方言 日本語諸方言コーパス モニター公開版	COJADS	✓	✓	コーパスの紹介ページ
日本語学習者 多言語母語の日本語学習者横断コーパス	I-JAS	✓	✓	I-JAS に関する詳細な情報は「I-JAS 関連資料」をご参照ください。プレインテキスト・音声・作文は「データ配布」からダウンロードできます。

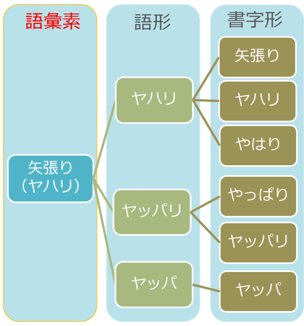
検索結果の例

検索語の品詞情報、検索語が現れる資料の情報を表示



形態論情報

語彙素という単語の見出し語に付与されている情報



形態論情報の見出し語は図のような階層構造を持っている

語彙素「矢張り」の例

語彙素：辞書の見出しのレベル
語形　：異語形を区別するレベル
書字形：異表記を区別するレベル

語彙素を使うことで表記の違いをまとめることができる

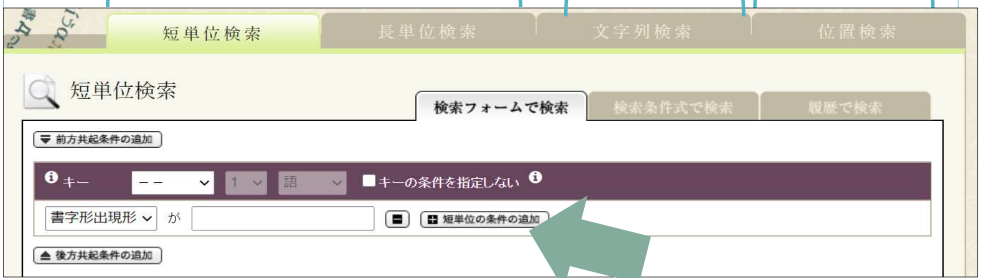
中納言の4つの検索方法

中納言では4種類の検索方法によって検索可能（一部コーパスでは長単位検索はない）

短単位検索・長単位検索
形態論情報を使った詳細検索

文字列検索
形態論情報を使わず出現形で検索

位置検索
出現位置から（再）検索



細かく情報を指定して検索可

文字列検索

単語単位を気にせず完全一致する検索が可能

表記の違い（読みや書き方の違い）は考慮しないため
その場合は短単位・長単位検索を利用する必要がある

短単位検索・長単位検索

形態論情報に基づいた検索ができる

短単位検索

短単位が一致する検索が可能

語彙素「日本」の検索例

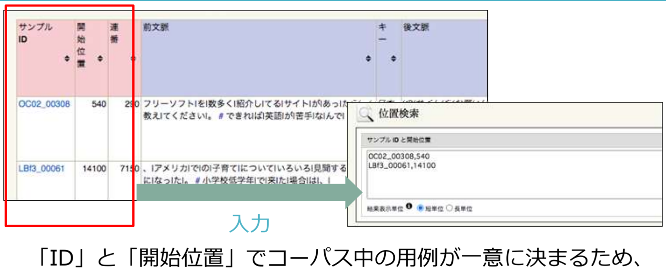
語彙素「日本」がヒットする
→　「日本」「ニッポン」「にっぽん」「にほん」
いずれも共通語彙素なのでヒット

長単位検索

長単位が一致する検索が可能

「日本語」「日本語研究」「日本語教育」などが長単位

位置検索



「ID」と「開始位置」でコーパス中の用例が一意に決まるため、
2つを保持していれば再現検索可能

中納言　https://chunagon.ninjal.ac.jp/