

筑波 ウェブコーパス

NINJAL-LWP for TWC (NLT)

ことばの特徴を知ろう！

11億語超の大規模コーパスを使って



11億語を超える

日本語のウェブサイトから集めた
言葉のデータ集



優しい
親切な

2語比較検索

2つの語を同時に比較して
語彙の共起関係を詳しく理解



こんにちは
你好
HELLO

多言語対応

日本語・英語・中国語(繁/簡)
メニューに対応

是非使ってネ！



WEB

<https://tsukubawebcorpus.jp/>

お問い合わせ

jp-kyoten@un.tsukuba.ac.jp

共同開発



筑波大学
University of Tsukuba



日本語・日本事情遠隔教育拠点
The Center for Distance Learning of Japanese and Japanese Issues

筑波ウェブコーパスとは？

Tsukuba Web Corpus : TWC

WEB <https://tsukubawebcorpus.jp/>

筑波大学が提供している「筑波ウェブコーパス」は、日本語のウェブサイトから収集された約11億3800万語のデータを使用して構築された大規模な日本語コーパスです。

専用の検索ツール **NINJAL-LWP for TWC(NLT)** を用いて、コーパス内の語彙の共起関係や使用頻度、文法的振る舞いを網羅的に調査・比較することができます。

早速
使ってみよう！



簡単に操作ができるので、

思考を途切らせることなく、調べることができます。

優しい 頻度=90,825

グループ別 **パターン頻度順** 基本

» 名詞

パターン	頻度	比率
優しい+名詞	37,064	<div style="width: 100%;"></div>
優しき+名詞	229	<div style="width: 1%;"></div>
優しい+接尾名詞	12,275	<div style="width: 15%;"></div>

名詞+助詞 «

パターン	頻度	比率
…が優しい	3,700	<div style="width: 4%;"></div>
…は優しい	2,797	<div style="width: 3%;"></div>
…も優しい	1,455	<div style="width: 1.6%;"></div>

専門的な
MI/LDスコアを使って
詳細な分析も！

優しい+名詞	3,738種類	頻度	MI	LD
コロケーション				
優しい人	1,674	4.02	4.10	
優しい言葉	1,231	6.53	6.50	
優しい気持ち	1,181	7.11	7.01	
優しい心	882	5.87	5.86	
優しいの	710	1.65	1.73	
優しいほう	598	3.78	3.83	
優しい声	534	6.24	6.11	
優しいもの	526	2.69	2.76	
優しい笑顔	522	8.21	7.55	

※検索画面のイメージ

前後のセンテンスを
表示するダイアログも
見られるよ！

優しい人 1,674件

- やさしい人です。
(今月の?スポットライト Bitter with the Sweets(ビターウィズサズイツ))
- 優しい人だった。
(第23回 鳴門教育大学国語教育学芸発表資料(平成20年))
- 優しい人なので。
(真面目な年下男性へのアプローチの仕方 | OKWave)
- 優しい人である。
(裁判傍聴記)

Page 1 / 17 100 1,674件中 1-100を表示

フィルター機能で
特定のコロケーションを
探すことも！

XLS/CSV形式で
ダウンロード可能！

2語比較検索



優しい+名詞 3,733

優しい

コロケーション	頻度	MI	LD
優しい人	1,674	4.02	4.10
優しい言葉	1,231	6.53	6.50
優しい気持ち	1,181	7.11	7.01
優しい心	882	5.87	5.86

親切な+名詞 712

親切な

コロケーション	頻度	MI	LD	LD差
親切な人	390	5.33	2.01	2.09
親切な言葉	24	4.25	0.93	5.57
親切な気持ち	15	4.22	0.88	6.13
親切な心	22	3.95	0.63	5.23

※検索画面のイメージ

詳しい使い方は
ユーザーズガイドを
見てね！

「優しい人」は1,674回出てくるけど
「親切な人」は390回しか出てこない！
↓
大まかな傾向を簡単に把握できる

簡単にソートして両方のコロケーションを詳細に分析！

多言語対応

筑波ウェブコーパスの専用検索ツール **NINJAL-LWP for TWC(NLT)** では日本語、英語、中国語(簡体字・繁体字)のメニューが利用可能です。
日本語だと使いづらいと感じていた海外の皆さまも是非ご利用ください！

利用者の声



実際の会話や文書に見られる自然な表現や文法の使用例を教室で活用することが出来た！

現代日本語の使用パターン、構文分析、意味論的研究が捗った！

単語や表現の使用頻度や文脈を調べることができ意味や使い方の理解を深めることができた！