

JEFL Corpus の使用法

研究生 尾崎ちひろ

目次

1. JEFLL Corpus とは
2. 主要な機能
 - 2.1 語句検索
 - 2.2 品詞検索
 - 2.3 共起検索
 - 2.4 単語リスト
3. 補助的機能
 - 3.1 表示設定
 - 3.2 サブコーパス
 - 3.3 ダウンロード
4. おわりに

1. JEFLL Corpus とは

- ・日本の中高生による自由英作文を約67万語集めたコーパス
- ・6種類のトピックで、20分間の辞書なしによる自由英作文
- ・オンライン上で無償公開
(<http://scn.jkn21.com/~jefll03/>)
- ・英語で表現できない場合は、日本語を使うことが許されている
 - 日本語表記は[JP:(ローマ字表記)]で表示

2. 主要な機能

2.1 語句検索

①表記形検索

- ・入力した語や句の形で検索

②基本形

- ・検索する動詞の全ての活用形が検索可能
- ・ソート機能、語(句)集計を使うと便利

③出現位置指定

- ・文頭または文末で使われている文を抽出できる

④2単語間の語数

- ・句動詞の抽出 例) take ~ out

JEFLL

ScN
Shogakukan
Corpus
Network

| 語(句)検索 | 品詞検索 | 共起検索 | 単語リスト | | 表示設定 | ソート | サブコーパス | ヘルプ |

take 基本形で検索 出現位置指定 なし 2単語間の語数 1 語以内

検索 クリア

①表記検索

②基本形

③出現位置設定

④2単語間の語数

ソート機能

ソート設定 | ヘルプ |

なし
 左ソート
 右ソート

詳細設定
 優先順位:

ソートキー

表記形
 品詞
 基本形
 大文字/小文字を区別

語(句)集計

JEFL Shogakukan Corpus Network
|閉じる|検索ウィンドウへ|
検索語(句): 基本形: take
総件数: 2179 (3255.73 / 1M語) 採用: 2179 サブコーパス指定: 無 << < 1 / 88 > >> | 1 | ページへ

1 00012	In the morning , I ca n't take	a lot of breakfast because I wake up so slow .
2 00291	So I take	money from bank .
3 00394	But I do n't take	a lot of money .
4 00394	[JP:sarani] I take	money [JP:no_hotondo_ha_chokin_shita] .
5 00545	Then I take	the train .
6 00661	I'm taking	the [JP:グローブ] and bat for 10 years old .
7 00668	I like takeing	very much .

語(句)集計

|ヘルプ|

集計範囲 ~ 頻度 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 含む 含まない

OK キャンセル 閉じる

2. 主要な機能

2.2 品詞検索

- ・綴りは同じだが複数の品詞を持つ単語を検索する際に、品詞の条件をつけて検索できる

2.3 共起検索

- ・コロケーションの使いこなしのチェックに役立つ機能

2.4 単語リスト

- ・品詞別使用頻度を表示

品詞検索

共起検索

品詞検索

語(句)検索 | ログアウト |

JEFL **ScN**
Shogakukan
Corpus
Network

| [語\(句\)検索](#) | [品詞検索](#) | [共起検索](#) | [単語リスト](#) | | [表示設定](#) | [ソート](#) | [サブコーパス](#) | [ヘルプ](#) |

take 基本形で検索 出現位置指定 なし ▾ 2単語間の語数 1 ▾ 語以内

品詞検索画面

品詞検索 | ログアウト |

JEFL **ScN**
Shogakukan
Corpus
Network

| [語\(句\)検索](#) | [品詞検索](#) | [共起検索](#) | [単語リスト](#) | | [表示設定](#) | [ソート](#) | [サブコーパス](#) | [ヘルプ](#) |

表記形: 基本形: 品 詞: [?] を 含む

共起検索画面

共起検索 | ログアウト |

JEFLL **ScN**
Shogakukan
Corpus
Network

| [語\(句\)検索](#) | [品詞検索](#) | **共起検索** | [単語リスト](#) | | [表示設定](#) | [ヘルプ](#) |

中心語: 基本形で検索 中心語の品詞: [\[?\]](#)

共起語の設定: リストを でソート 品詞設定: [\[?\]](#)

品詞検索画面

単語リスト | ログアウト |

JEFLL **ScN**
Shogakukan
Corpus
Network

| [語\(句\)検索](#) | **品詞検索** | [共起検索](#) | [単語リスト](#) | | [ヘルプ](#) |

単語: 基本形で検索

品詞: [\[?\]](#) を 表示件数: / ページ

単語リストの簡単な使い方と注意

3. 補助的機能

3.1 表示設定

- ・検索結果の表示方法が変更可能

3.2 サブコーパス

- ・特定のサブコーパスを選んで検索可能

3.3 ダウンロード

- ・検索結果をPCに保存し、データを加工することが可能

表示設定

サブコーパス

語(句)検索

| ログアウト |

JEFLL

SCN
Shogakukan
Corpus
Network

| 語(句)検索 | 品詞検索 | 共起検索 | 単語リスト | | 表示設定 | ソート | サブコーパス | ヘルプ |

take

基本形で検索 出現位置指定 なし 2単語間の語数 1 語以内

検索

クリア

表示設定画面

表示設定

| ヘルプ |

検索結果上限 3000 件 1ページ行数 25 行

レイアウト 1行 120 文字 1文KWIC表示 1文表示

属性表示 表記形 品詞 基本形

OK

キャンセル

閉じる

サブコーパス画面

JEFLC サブコーパス			
保存しないで閉じる		設定を保存して閉じる	
		全てを解除	語数の再カウント
ヘルプ			
<input checked="" type="checkbox"/> ALL	genre	<input type="checkbox"/> argumentative <input type="checkbox"/> narrative	リセット
	year	<input type="checkbox"/> j1 <input type="checkbox"/> j2 <input type="checkbox"/> j3 <input type="checkbox"/> s1 <input type="checkbox"/> s2 <input type="checkbox"/> s3	リセット
	topic	<input type="checkbox"/> breakfast <input type="checkbox"/> dream <input type="checkbox"/> earthquake <input type="checkbox"/> festival <input type="checkbox"/> otoshidama <input type="checkbox"/> urashima	リセット
	school level	<input type="checkbox"/> high <input type="checkbox"/> mid <input type="checkbox"/> low	リセット
	school	<input type="checkbox"/> national <input type="checkbox"/> public <input type="checkbox"/> private	リセット
ファイル数 10038 延べ語数 669281 (比率 100%) 異なり語数 17464			

ダウンロード

JEFLC		Shogakukan Corpus Network	
閉じる 検索ウィンドウへ		ソート 語(句)集計 表示変更 ダウンロード ヘルプ	
検索語(句): 基本形: take			
総件数: 2179 (3255.73 / 1M語)		採用: 2179	サブコーパス指定: 無
		<< < 1 / 88 > >>	1 ページへ
1 00012	In the morning , I ca n't take	a lot of breakfast because I wake up so slow .	
2 00291	So I take	money from bank .	
3 00394	But I do n't take	a lot of money .	
4 00394	[JP:sarani] I take	money [JP:no_hotondo_ha_chokin_shita] .	
5 00545	Then I take	the train .	
6 00661	I'm taking	the [JP:グローブ] and bat for 10 years old .	
7 00668	I like takeing	very much .	

4. おわりに

JEFLC Corpusを用いると、中学生・高校生がしがちなエラーが重点的に指導すべきエラーか、軽く指摘するべきなのか、あるいはそのエラーは無視すべきなのか判断の材料に使える