

# 日本語指標比喩のデータセット の紹介とその活用

菊地 礼

国立国語研究所 [非]

- 「『現代日本語書き言葉均衡コーパス』に基づく指標  
比喩データベース」 (加藤ほか2020) について
  - どのような事例を収録しているか
  - どのように事例を集めたか
  - どのような情報がアノテーションされているか
- 本データの活用事例
  - 直喩を構成する意味分野が明らかにできたりする！

# 目次

- 1 : はじめに
- 2 : データの概要
- 3 : データの活用事例
- 4 : おわりに

紹介するデータ  
指標比喩とは  
比喩とデータ

## • 紹介するデータ

- 「『現代日本語書き言葉均衡コーパス』に基づく指標  
比喩データベース」 (**BCCWJ-simile**)
  - コーパスから指標比喩を収集し、アノテーション
- BCCWJコアデータ6レジスタ
  - Yahoo!知恵袋 白書 Yahoo!ブログ 書籍 雑誌 新聞
  - 1,290,060語57,256文
- 全822例収録
- 「BCCWJデータ配布サイト」にてDL可能

| 1 | サンプルID        | 前文脈   | 比喻指標 | 後文脈  | 比喻表現 (該当部)  | 比喻表現 (該当部) 整理   | 被喩辞            | 喩辞   | 類型A            | 付属語A | 類型B  | 付属語E | 類型C |
|---|---------------|---|------|--|---|---|----------------|------|----------------|------|------|------|-----|
| 2 | OC01_00004-17 | スーパーファミコンで1番最初にでたソフトって何ですか??#やっぱ、マリオワールドですかね??#じみにアクトレイザーも早かったですよね。#1:FZER0#発売日:九十/十一/二十一#ジャンル:レース 価格:七千円#内容説明:近未来レーシングゲーム#2:スーパーマリオワールド#発売日:九十/十一/二十一#ジャンル:アクション 価格:八千円#内容説明:恐竜ランドを舞台(に) | し    | たスーパーマリオ   | 恐竜ランドを舞台(に)***たスーパーマリオ  | 恐竜ランドを舞台にしたスーパーマリオ  | #話の中心になる場所     | 舞台   | 恐竜ランド          | ヲ    | 舞台   | ニ    | した  |
| 3 | OC01_01579-50 | 。腰振ったらダメだろ……。古い話ですが、この木なんの木気になる木?♪のCMの大きな木はどこにあるのですか?ハワイ・オアフ島(アメリカ)にあるそうですよ。参考URLにそのCM関連のサイトを掲載しておきました。ご覧ください。北海道札幌市上空に飛行機雲でPOCARISWEATとかいてあった!まさしくポカリのCMだよ!!!!                           | 本当   | に飛行機雲のようにして空に書いているんですよ・・・・・・・・・・・・・・・・カネボウのCMで男の子の胸に炎の穴を開ける少し桃井かおり似の女性の詳細を御存知の方、教えてください。似てないような気がするんですけど・・・・・・・・麻生久美子さんです。。理由はわかりませんがうちではニッポン放送がとても聞きづらいです | 北海道札幌市上空に飛行機雲でPOCARISWEATとかいてあった!まさしくポカリのCMだよ!!!!**に飛行機雲のようにして空に書いているんですよ | 北海道札幌市上空に飛行機雲でPOCARISWEATとかいてあった!まさしくポカリのCMだよ!!!!**に飛行機雲のようにして空に書いているんですよ | #POCARISWEATの字 | 飛行機雲 | #POCARISWEATの字 | ハ    | 飛行機雲 | ノヨウ  |     |

## • 指標比喩とは

### • 比喩の分類 (国立国語研究所1977)

#### • 指標比喩：形式的な有標性を持つ

• e.x. 「雪のような肌」

#### • 結合比喩：語と語の意味的逸脱によって成立する

• e.x. 「彼女はバラだ」

#### • 文脈比喩：文脈における意外性によって成立する

• e.x. 「峠を越した」

## • 比喩とデータ

### • 作例・内省の研究がメイン

#### • もしくは特定のテキストから収集

• たとえば → 芥川龍之介の比喩 綿矢りさの比喩…

### • 比喩は用例の収集が困難

#### • 出現頻度が低い

#### • 特定の形式を持つことが少ない

• 指標比喩は特徴的な形式を有する

• 手掛かりにして、データを構築しよう！

## • 比喩とデータ（続）

### • 比喩の収集とアノテーションの進行

#### • 現代語コーパス：Kato et.al. (2022)

- 結合比喩・文脈比喩なども含めて

#### • 古典語コーパス：菊地 (2023)

- 指標比喩・結合比喩・文脈比喩など

#### • 青空文庫：Komatsubara (2021)

- レトリック全般

言語資源に  
基づいた研究の基礎



## • ちなみに…



### ChatGPT

直喩の実例をいくつか挙げます。これらの例は、異なる文脈やシチュエーションで使われることがあります。いずれも一つの事物を別の事物に直接的に等しいと表現することで、意味やイメージを強化しています。

1. 彼は獅子だ。
2. 彼女は太陽だ。
3. その子は風の子だ。
4. 彼の心は鋼のようだ。
5. 彼女の目は星だ。
6. その場所は天国だ。
7. 彼の声は雷鳴だ。
8. 彼女は花だ。
9. その日は災害だ。
10. その子供は天使だ。

隠喩と区別  
できていない

生成AIは  
レトリックが苦手？

規範から外れる言語  
の研究も進める必要

# 目次

- 1 : はじめに
- 2 : データの概要
- 3 : データの活用事例
- 4 : おわりに

指標比喩の収集方法  
アノテーション

## • 指標比喩の収集方法

- ①：手掛かり句の認定と検索
  - 比喩指標要素：「まるで」「よう」など
  - 指標要素近傍の前後100字を読む
- ②：MIPによる比喩的語義の認定 ( Pragglejaz Group2007, Steen et al.2020 )
  - 基本義・代表義の認定 (宮島1972)
  - 基本義・代表義から逸脱するものを比喩と認定
- ③：アノテーション
  - 喩辞・被喩辞の同定、結合の抽出、結合の種別など

## • アノテーション

### • 形式関連の情報

- **被喻辞** : 喩えられる対象
- **喩辞** : 喩える事物・事柄
- **比喩的結合の類型** : 要素の結びつき

| 前文脈  | 比喩指標 | 後文脈  | 比喩表現 (該当部)   | 比喩表現 (該当部) 整理   | 被喻辞    | 喩辞 | 類型A    | 付属語A | 類型B | 付属語B | 類型C  | 付属語C |
|--|------|--|--|---|--------|----|--------|------|-----|------|------|------|
| の燃料高騰の煽りを受けて集魚灯を使うイカ漁が出漁の危機にあってます。昨年から比べて数倍の燃料の値上がりに赤字覚悟のイカ漁と漁師さんは嘆いています。北海道の夏の味覚、イカソーメン、イカ刺し、イカの一 | に    | 細く切り、好みで生姜醤油や山葵醤油、めんつゆにつけて食べる料理です?♪ 鮮度の良いものは透き通っていて、歯ごたえと甘みを堪能することができます^^ 別の古道さ入る 林を通り、途中で立ち止まって山を見る 真ん中あたりに見える綺麗な三角山の左隣 | イカソーメンは、生のイカを二枚あるいは三枚に分け、素麺のよう***細く切り、好みで生姜醤油や山葵醤油、めんつゆにつけて食べる料理です?♪ | イカソーメンは、生のイカを二枚あるいは三枚に分け、素麺のよう**細く切り、好みで生姜醤油や山葵醤油、めんつゆにつけて食べる料理です?♪ | イカソーメン | 素麺 | イカソーメン | ハ    | 素麺  | ノヨウニ | 細く切る |      |

## • アノテーション (続)

### • 意味関連の情報

- **分類番号** : 『分類語彙表』の意味番号
- **結合種別** : 要素の結合の仕方、概念マッピングなど

| タイプA<br>▼  | 付属<br>語A<br>▼ | タイプB<br>▼  | 付属<br>語B<br>▼ | タイプC<br>▼ | 付属<br>語C<br>▼ | 結合種別<br>▼ | 結合追加<br>▼          |
|------------|---------------|------------|---------------|-----------|---------------|-----------|--------------------|
| 1.550<br>6 | ガ             | 1.431<br>0 |               |           |               | 転換        | 1.4310 ノヨウニ 3.1912 |

## • アノテーション (続続)

### • 印象評定データ

• 以下の項目を一般日本語話者によって 0～5 で評価

• 比喩性 新奇性 わかりやすさ 擬人化 具体化

| sent   | metaphor | new   | understar | prosopop | concrete |
|--|----------|-------|-----------|----------|----------|
| イカソーメンは、生のイカを二枚あるいは三枚に分け、素麺のように細く切り、好んで生姜醤油や山葵醤油、めんつゆにつけて食べる料理です?♪ | 2.800    | 1.200 | 4.067     | 1.300    | 3.233    |

# 目次

- 1 : はじめに
- 2 : データの概要
- 3 : データの活用事例
- 4 : おわりに

本データの活用事例  
直喩を構成する語の意味分野とは  
『分類語彙表』の意味番号  
直喩を構成する語の意味分野（体言）  
直喩を構成する語の意味分野（用言）

## • 本データの活用事例

- 直喩を構成する語の意味分野：本節
- 希少な形式の確保：菊地（2020）
  - 「～と思う」「～と考える」などを用いた直喩の成立条件
- 印象評定への着目：加藤ほか（2020）
  - 人に関わる例（擬人化など）は評定値が高くなる傾向
  - 身体性を基盤とした比喻理解



- 直喩を構成する語の意味分野とは

- 「雪のような肌」

- 「雪」：自然物

- 「肌」：人間の身体

- 『分類語彙表』の意味番号

- 「単語が表わし得る意味の世界を分類して、  
その分類の各項にそれぞれの単語を配当したもの」

(『分類語彙表増補改訂版』 p.9)

## 『分類語彙表』の意味番号

### 類

体の類(1)  
用の類(2)  
相の類(3)  
その他(4)

### 部門・中項目

抽象的關係(10~19)  
人間活動の主体(20~27)  
人間活動-精神・行為(30~38)  
生産物・用具(40~47)  
自然物・自然現象(50~57)

### 分類項目

こそあど・他  
真偽・是非  
本体・代理  
など

1.4550 (体-生産物・用具-道具-刃物)

# データの活用事例 4/7

|           |    |           |     |            |    |             |      |
|-----------|----|-----------|-----|------------|----|-------------|------|
| 1.10 (事柄) | 20 | 1.21 (家族) | 13  | 1.34 (行為)  | 29 | 1.46 (機械)   | 39   |
| 1.11 (類)  | 23 | 1.22 (仲間) | 6   | 1.35 (交わり) | 20 | 1.47 (土地利用) | 22   |
| 1.12 (存在) | 5  | 1.23 (人物) | 19  | 1.36 (待遇)  | 8  | 1.50 (自然)   | 31   |
| 1.13 (様相) | 32 | 1.24 (成員) | 29  | 1.37 (経済)  | 17 | 1.51 (物質)   | 59   |
| 1.14 (力)  | 8  | 1.25 (公私) | 24  | 1.38 (事業)  | 7  | 1.52 (天地)   | 47   |
| 1.15 (作用) | 35 | 1.26 (社会) | 38  | 1.40 (物品)  | 28 | 1.53 (生物)   | 3    |
| 1.16 (時間) | 25 | 1.27 (機関) | 9   | 1.41 (資材)  | 25 | 1.54 (植物)   | 29   |
| 1.17 (空間) | 38 | 1.30 (心)  | 118 | 1.42 (衣料)  | 20 | 1.55 (動物)   | 32   |
| 1.18 (形)  | 33 | 1.31 (言語) | 61  | 1.43 (食料)  | 33 | 1.56 (身体)   | 135  |
| 1.19 (量)  | 15 | 1.32 (芸術) | 40  | 1.44 (住居)  | 53 | 1.57 (生命)   | 17   |
| 1.20 (人間) | 87 | 1.33 (生活) | 44  | 1.45 (道具)  | 49 | 総数          | 1425 |

## • 直喩を構成する語の意味分野（体言）

- 1.20 (人間)
- 1.56 (身体)
- 1.44 (住居)
- 1.51 (物質)
- 1.30 (心)
- 1.31 (言語)

ヒト関連

モノ関連

抽象物

視覚的な類似性に傾く (加藤2018)

抽象物の具体化

# データの活用事例 6/7

|           |     |            |    |           |     |
|-----------|-----|------------|----|-----------|-----|
| 2.10 (真偽) | 0   | 2.19 (量)   | 0  | 2.37 (経済) | 3   |
| 2.11 (類)  | 0   | 2.30 (心)   | 28 | 2.38 (事業) | 9   |
| 2.12 (存在) | 20  | 2.31 (言語)  | 14 | 2.50 (自然) | 2   |
| 2.13 (様相) | 3   | 2.32 (芸術)  | 5  | 2.51 (物質) | 7   |
| 2.14 (力)  | 0   | 2.33 (生活)  | 18 | 2.52 (天地) | 0   |
| 2.15 (作用) | 103 | 2.34 (行為)  | 78 | 2.56 (身体) | 0   |
| 2.16 (時間) | 0   | 2.35 (交わり) | 2  | 2.57 (生命) | 2   |
| 2.17 (空間) | 1   | 2.36 (待遇)  | 9  | 総数        | 304 |

## • 直喩を構成する語の意味分野 (用言)

- 2.15 (作用)
  - 2.34 (行為)
  - 2.30 (心)
- 物理的  
動作・活動
- 心理が  
関わる動作
- 動きがどのように  
なされたか  
(視覚の具体化)
- 鳩をタオルで包むように抱きかかえて〔後略〕 (OY05\_01367-3480)
  - 狂ったように笑いながらデジカメのシャッターを押し続けた。  
(PM25\_00084-78060)

# 目次

- 1 : はじめに
- 2 : データの概要
- 3 : データの活用事例
- 4 : おわりに

本発表のまとめ  
今後の展望

## • 本発表のまとめ

### • BCCWJ-simileについて…

- ①： 比喩指標要素を手掛かりとして、MIPに基づく基本義からの逸脱事例を収集した
- ②： アノテーション
  - 形式： 喩辞 被喩辞 結合類型
  - 意味： 『分類語彙表』 分類番号 結合種別
  - 印象評定： 比喩性 新奇性 わかりやすさ 擬人化 具体化
- ③： 直喩に用いる意味分野を確認するなどの活用例の紹介
  - 視覚への偏りの確認



## • 今後の展望

- BCCWJ-Metaphor (Kato et al.2022) の構築
  - 結合比喩その他との比較が可能
- 古典語比喩データ (菊地2023)
  - 歴史的な比較が可能
- 日本語レトリックコーパス (Komatsubara2021)
  - 他のレトリックとの比較が可能

資源が整備されることで  
多様な手法・観点で研究  
が遂行可能

- 加藤祥（2018）「隠喩と直喩の違いは何か：用例に見る隠喩と直喩の使い分けから」『認知言語学研究』3巻、pp.1-22：開拓社.
- 加藤祥・菊地礼・浅原正幸（2020）「『現代日本語書き言葉均衡コーパス』に基づく指標比喩データベース」『自然言語処理』27巻4号、pp.853-887：言語処理学会.
- Kato Sachi, Kikuchi Rei, Asahara Masayuki（2022）“Figurative Expression Information Database on `Balanced Corpus of Contemporary Written Japanese’”, 15th Researching and Applying Metaphor Conference, abstract.
- 菊地礼（2023）「中世期日本語比喩表現の収集の試み」『言語資源ワークショップ発表論文集』1号、pp.224-234：国立国語研究所.
- 国立国語研究所（1977）『比喩表現の理論と分類』国立国語研究所報告57：秀英社.
- 国立国語研究所編（2004）『分類語彙表増補改訂版』大日本図書.

- Komatsubara Tetsuta, 2021, "The Corpus of Japanese Figurative Language : Toward a comprehensive framework for describing figurative language", 『国際文化研究：神戸大学大学院国際文化学研究所紀要』 55号、pp.107-134：神戸大学大学院国際文化学研究所。
- 宮島達夫（1972）『動詞の意味・用法の記述的研究』秀英出版。
- Pragglez Group（2007）"MIP: A Method for Identifying Metaphorically Used Words in Discourse." *Metaphor and Symbol* 22（1）, pp.53-62.
- Steen, G. J., Dorst, A. G., Herrman, J. B., Kaal, A., Krennmayr, T., and Pasma, T.（2010）. *A Method for Linguistic Metaphor Identification*. John Benjamins Publishing.

- 本発表はJSPS22K18483の助成を受けたものである。