

## 『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 祥<sup>†</sup>・菊地 礼<sup>††</sup>・浅原 正幸<sup>†††</sup>

日本語の比喩表現の実態把握を目的として、『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を構築した。『比喩表現の理論と分類』に掲載されている359種類の比喩指標要素を手掛かりとし、『分類語彙表』に基づいて類義用例を確認しながら指標比喩表現候補を展開し、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語から人手で822件抽出した。抽出した比喩用例には、喩辞・被喩辞の情報と、その分類語彙表番号を付与したほか、擬人化・擬物化・擬生化・具象化などの種別情報も付与した。さらに提喩・換喩・文脈比喩・慣用表現などの情報も付与した。上記作業は言語学者によったが、非専門家が比喩表現をどのように捉えるかを評価するために、比喩性・新奇性・わかりやすさ・擬人化・具体化（具象化）の5つの観点について、1事例あたり22-77人分（平均33人分）の評定値を付与した。レジスタ毎の相対度数や評定値の分布により、現代日本語の指標比喩表現の使用傾向を確認した。

キーワード：直喩・比喩表現・比喩指標・語義タグ付きコーパス

## Database of Figurative Expressions with Indicators from the ‘Balanced Corpus of Contemporary Written Japanese’

SACHI KATO<sup>†</sup>, REI KIKUCHI<sup>††</sup> and MASAYUKI ASAHARA<sup>†††</sup>

A figurative expression database was constructed based on the Balanced Corpus of Contemporary Written Japanese (BCCWJ), with the goal of understanding actual usage of figurative expressions in Japanese. Using the three hundred fifty nine types of figurative expression indicators listed in ‘A Stylistic Study of the Figurative’ (*Hiyuhyogen-no Riron-to Bunrui*) as clues for metaphor indicator elements, candidates were selected based on synonym examples confirmed in the ‘Word List by Semantic Principles’, and a total of eight hundred twenty two expressions were manually extracted from one million two hundred ninety thousand sixty words found in six registers of core data (Yahoo! Answers, white papers; Yahoo! Blog, books, magazines, and newspapers). In addition to the vehicle, topic, and Word List by Semantic Principles label of each metaphor example, type categories such as personification, objectifi-

---

<sup>†</sup> 目白大学, Mejiro University

<sup>††</sup> 中央大学, Chuo University

<sup>†††</sup> 人間文化研究機構 国立国語研究所, NINJAL, Japan

cation, biomimicry, and substantiation were defined. Examples were also classified into categories such as synecdoche, metonymy, contextual metaphor, and idiomatic expression. Although the work above was carried out by linguists, ratings were also assigned to each example for five aspects (figurativeness, novelty, comprehensibility, personification, and substantiation) based on evaluations by twenty two to seventy seven non-experts (average: thirty three) to evaluate how these figurative expressions were perceived. The usage trends for each of these figurative expression indicators in contemporary Japanese were determined based on their relative frequency in each register and distribution of their rating values.

**Key Words:** *Simile, Figurative Expression, Metaphor Indicator, Sense Tagged Corpus*

## 1 はじめに

比喩表現は、意味解釈の構成性の要請を満たさない事例の代表である。Lakoff and Johnson (1980) (日本語訳 (レイコフ, ジョンソン 1986)) は「思考過程の大部分が比喩によって成り立つ」と言及している。言語学においても、そもそも形態や語彙、辞書構造、文法をはじめ、言語の大部分が比喩的な性質に基づくこととされ、比喩研究は「言語の伝達のメカニズムを理解していくための基礎的な研究」と位置づけられる (山梨 1988)。また、言語処理においても基本義からの転換という現象が意味処理の技術的障壁になっている。比喩表現データベースは、言語学・言語処理の双方で求められている重要な言語資源である。そこで我々は、『現代日本語書き言葉コーパス』(Maekawa, Yamazaki, Ogiso, Maruyama, Ogura, Kashino, Koiso, Yamaguchi, Tanaka, and Den 2014) (以下 BCCWJ と呼ぶ) コアデータ 1,290,060 語 57,256 文に基づく大規模比喩表現データベースを構築した。

比喩性の判断は、受容主体の主観によるものであり、形式意味論的な妥当性・健全性を保持しうるものではない。我々は、研究対象となる比喩表現が適切に含まれるような作業手順として MIP (Metaphor Identification Procedure) (Pragglejaz Group 2007) を拡張した MIPVU (Metaphor Identification Procedure VU University Amsterdam) (Steen, Dorst, Herrmann, Kaal, Krennmayr, and Pasma 2010) を取り入れる。さらに、安定的に一貫して抽出するため、先行研究の中でもより形式的に比喩を捉える中村 (1977) の研究に倣い、喩辞 (喩える表現) の基本義からの語義の転換・逸脱と喩辞に関連する要素の結合に着目する。喩辞の語義の転換・逸脱の判断には、『分類語彙表』(国立国語研究所 2004) に基づいた語義を用い、被喩辞 (喩えられる表現) との語義の差異を検討する<sup>1</sup>。さらに、被喩辞相当の語義があるべき箇所に喩辞の語義が現れる表現中の要素の結合における比喩的な転換・逸脱の有無を確認する。比喩表現と考えられる部分について、喩

<sup>1</sup> 本稿では、喩える表現・語を「喩辞」、喩えられる表現・語を「被喩辞」と呼ぶ。それぞれ、「喩詞」と「被喩詞」、「ソース (source)」と「ターゲット (target)」、「サキ」と「モト」、「媒体 (vehicle)」と「主題 (topic)」と呼ばれるものに相当する。

辞相当の出現箇所を同定するとともに、比喩関連情報をアノテーションする。但し、非専門家が比喩表現と認識しない表現を多く含む結果となるため、非専門家の判断としてクラウドソーシングによる比喩性の判断を収集する。

我々が構築した指標比喩データベースは、以下のもので構成される：

- 比喩表現該当部 (3.4 節)
- 比喩指標要素とその類型 (中村 1977) (3.2 節), その分類語彙表番号 (3.2 節, 3.3 節)
- 比喩的転換に関わる要素の結合とその類型 (3.2 節), その分類語彙表番号 (3.3 節)
- 概念マッピングにおける喩辞・被喩辞 (3.5 節), その分類語彙表番号 (3.3 節)
- 概念マッピングに基づく比喩種別 (擬人・擬生など) (3.5 節)
- 非専門家の評定値 (比喩性・新奇性・わかりやすさ・擬人化・具体化) (3.6 節)。

本稿では、そのデータ整備作業の概要を示すとともに構築したデータベースの基礎統計や用例を示す。本研究の貢献は次の通りである。まず、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 文に基づく、日本語の大規模指標比喩データベースを構築した。この指標比喩データベース構築において、まず英語で実施された比喩用例収集手法である MIP, MIPVU に対して『分類語彙表』の語義に基づく手法を提案し、日本語の比喩用例収集作業手順を整理した。本作業に必要な比喩用例収集の手掛かりとなる中村 (1977) の比喩指標要素 (359 種類) を電子化し、新たに分類語彙表番号を付与し、再利用可能な比喩指標要素データベースを整備した。また、収集した比喩表現に対し、喩辞・被喩辞・分類語彙表番号・比喩種別などをアノテーションした。さらに、収集した指標比喩を刺激としてクラウドソーシングによる質問紙調査を実施し、非専門家の比喩性判断を収集した。構築した大規模指標比喩データベースに基づく調査が可能となったため、比喩表現の遍在性を確認し、非専門家の比喩性判断の実態を明らかにした。

本稿の構成は次のとおりである。2 節に関連研究を示す。3 節ではデータ整備の概要について解説する。4 節ではデータの集計を行い、指標比喩の分布を概観する。5 節にまとめと今後の方向性について示す。

## 2 関連研究

以下では、比喩表現コーパス・比喩表現用例集の先行研究について確認する。2.1 節では、英語に関する先行研究として Amsterdam Metaphor Corpus (Steen et al. 2010) とその作業手順である Metaphor Identification Procedure (MIP) (Pragglejaz Group 2007) について解説する。2.2 節では、我々の研究の基底となる中村 (1977) の『比喩表現理論と分類』における日本語の比喩表現の定義について示す。2.3 節では、その他の日本語の比喩表現コーパスおよび比喩表現用例集について示す。

## 2.1 Amsterdam Metaphor Corpus とその作業手順 MIP

英語における大規模な比喩コーパスの先行研究として、Amsterdam Metaphor Corpus (Steen et al. 2010) がある。BNC-Baby コーパスから抽出した4レジスタ約19万語を用い、隠喩 (indirect) と直喩 (direct)、暗黙 (implicit) の3種類の比喩の関連語を検索可能にしたもので、MIP (Metaphor Identification Procedure) (Pragglejaz Group 2007) を拡張したMIPVU (Metaphor Identification Procedure VU University Amsterdam)<sup>2</sup> (Steen et al. 2010) により、語単位の比喩性が判定され、直喩の指標 (metaphor signals) と概念マッピング (conceptual mappings) 情報が付与されている。

図1にMIPの概要(著者訳) (Pragglejaz Group 2007) を示す。テキストを読んで、テキストの文脈内における語義を判定し、基本義からの転換を認識する作業手順であるが、語義の判定基準および基本義の定義などは言語依存であろう。3.1節では、このMIPに基づく日本語の比喩表現収集手順の概要を図1に対応する形で示す。

MIPVUはMIPが認定する比喩表現を直喩や境界事例にも拡張したものである。図2にMIPVUによるMIPの拡張(著者訳) (Steen et al. 2010) を示す。比喩表現と見なす収集対象として、隠喩のみならず、直喩や比喩の境界事例を含む。語義の典拠として、主にMacmillan Dictionary<sup>3</sup>を用い、Macmillan Dictionaryでは判断できない場合に、Longman Dictionary of

- (1) テキストの文脈全体を読み、意味の一般的な解釈を規定する。
- (2) テキストの文脈における語彙単位を決定する。
- (3)
  - (a) テキストの各語彙単位について、文脈におけるその意味、つまり、テキストに喚起されるシチュエーション(前後関係の意味)でそれが実体、関係、または属性にどのように適用されるかを規定する。語彙単位の前後にくるものを考慮する。
  - (b) 各語彙単位について、他の文脈では、所定の文脈に比べてより基本的な最新の意味を持つか否かを決定する。我々の目的上、基本的意味の傾向は
    - より具体的で、それらが喚起するものは、想像しやすく、見やすく、聞きやすく、感じ取りやすく、嗅ぎ取りやすく、味わいやすいものである。
    - 身体的な動作に関連がある。
    - より厳密(漠然としたものとは対照的に)である。
    - 歴史的に古い。
 基本的意味が、必ずしも語彙単位で最も頻出する意味とは限らない。
  - (c) 語彙単位が他の文脈で所定の文脈に比べてより基本的な最新の意味を持つ場合、前後関係の意味は基本的意味と対照的でありながら、より理解されうるものなのか否かを決定する。
- (4) そうであれば、その語彙単位を比喩的なものとしてマークする。

図1 Metaphor Identification Procedure (MIP) の概要(著者訳)

<sup>2</sup> <http://www.vismet.org/metcor/documentation/MIPVU.html>

<sup>3</sup> <https://www.macmillandictionary.com/>

- (比喩表現は) (品詞などの) 単語のクラスの境界は超えられない。例えば, 動詞の前後関係の意味は, 名詞としてのその基本的意味と同等と見ない
- (その上で比喩表現としての抽出対象は) 次のものを含む:
  - － 直喩
  - － 隠喩
  - － 比喩の境界事例
  - － 比喩のシグナル (指標・結合など)
  - － 擬人化のための比喩
- (比喩表現の認定にあたって) 単語の由来はまれな例でのみ考慮される

図 2 MIPVU による MIP の拡張 (著者訳)

Contemporary English<sup>4</sup>を用いる。2つの辞書でも判断がつかない場合は, Oxford English Dictionary<sup>5</sup>を用いる。基本義の認定に, 通時的な語源を重視しないことが明記されている。また, 比喩と判断されうるものを網羅的に抽出することを目指しているが, 概念メタファー (conceptual metaphor) は認定しない。

## 2.2 中村の『比喩表現の理論と分類』

日本語比喩表現の実態調査に, 明治～昭和期の文学作品 50 種から比喩情報の抽出を行った国立国語研究所報告 57『比喩表現の理論と分類』(中村 1977)があり, 我々の研究の比喩の定義は中村 (1977)の研究に基づく。中村は, 日本語の比喩表現を, 受容過程における比喩性把握に着目し, 次の3類型に分類している:

- 第1類 指標比喩  
受容主体が表現主体の比喩意識を感じ取る, 換言すれば, 受け手側での比喩の成立に直接に形式的に慣用している, 特定の言語形式をそなえており, それを<中略>2項間の関係の特異性としてでなく, 他から独立に抽出できる種類の比喩表現
- 第2類 結合比喩  
何らかの言語単位の結びつきに, 慣用からの顕著な逸脱または非論理性, 少なくとも言語上の論理的な飛躍が感じられる種類の比喩表現
- 第3類 文脈比喩  
比喩の目印となる指標も, 要素間の結合上の異常性も特に認められないが, その表現形式が表す言語的な意味と, それがその場であらわしていると思われる個別的な意味との対応に慣用から著しいずれが意識される種類の比喩表現

第1類の指標比喩は, 手掛かり句を持つ表現で技法としては直喩 (simile) に相当し, 中村は

<sup>4</sup> <https://www.ldoceonline.com/>

<sup>5</sup> <https://www.oed.com/>

例文 (1) をあげている。ここでは、実際に行われた事柄の非現実性を表現するために、「宙を歩いている」という表現により喩えている<sup>6</sup>とするが、「ような気がする」という形式によって、比喩であることが把握される例である。

(1) 耕はひたいに汗をおほえた。息が苦しくなった。宙を歩いているような気がする。

(「顔」丹羽文雄)

第2類の結合比喩は、技法としての隠喩 (metaphor) に相当し、中村は例文 (2) をあげている。例文中には複数の比喩表現がみられ、仮想の抽象体である「えたいのしれない不吉な塊」が、精神活動の抽象体である「私の心」に対して、物理的な働きかけである「圧えつけ」を行っている。全体として文字どおりの解釈ができない見立てとする。

(2) えたいのしれない不吉な塊が私の心を始終圧えつけていた。

(「檸檬」梶井基次郎)

第3類の文脈比喩は、言語形式も構成要素間の結びつきの異常も持たない表現である。技法としては諷喩 (allegory) にあたり、中村は例文 (3) をあげている。「打つべき釘を、打ち残した気持ち」は、大工仕事をしたあとの気持ちを表すことも可能であり、「気組みを守った。」まで読んだ全体としても文字どおりの解釈が可能である。しかし、前後の文脈から文字どおりの意味ではなく何かの見立てであると感じる場合、文脈中において文脈との無縁性のために比喩となる例である。

(3) 打つべき釘を、打ち残した気持ちがつづき、香奈江は、答えをせぬことで、わずかに相手をはねかえす気組みを守った。

(「風ふたたび」永井龍男)

中村は「表現主体の比喩意識の反映とみられる何らかの言語形式をそなえており、それが受容主体の比喩把握に一役買っている」場合を「A型把握」とし、A型把握によって収集された指標比喩にあたる用例を1,617種類の類型に整理し、359種類の比喩指標要素を体系化した。本研究においては、この中村の比喩指標要素の近傍に出現する比喩表現の収集を行う。指標比喩の全数収集を目指す。結合比喩・文脈比喩が認められた場合にはデータベースに登録する。作業に先立ち、同書に含まれる比喩指標要素などの電子化を行った。詳細については3.2節に示す。

## 2.3 その他の日本語の比喩表現用例集および比喩表現コーパス

本節では、中村 (1977) 以外の日本語の比喩表現用例集および比喩表現コーパスの先行研究について紹介する。最初に指標比喩 (いわゆる直喩) 関連の研究を紹介し、次に結合比喩 (いわ

<sup>6</sup> 比喩を表す場合に「喩える」を用い、例示を表す場合に「例える」を用いる。クラウドソーシング調査における実験協力者への教示は、「たとえる」を用いた。

ゆる隠喩) 関連の研究を紹介する。最後に書籍として出版されている比喩表現用例集を示す。

**指標比喩関連:** 比喩表現の収集にあたり, 比喩であることを明示する, 比喩の指標 (手掛かり句) となる表現を含む指標比喩 (いわゆる直喩) は, 指標を用いた比較的容易な収集が期待される。過去の研究は比喩指標要素「ようだ」のみを対象とするものが多い。これらすべての比喩指標要素を含む表現を収集すれば, 多くの直喩用例が収集可能であるように思われる。しかし, 指標は必ずしも比喩表現のみに含まれる語句ではない。指標を手掛かりにコーパスから用例収集を試みる際にも, 人手による判別作業が必須である。また, 比喩指標の語句は, 類語や言い換えの可能な表現があり得るため多様であり, 類似表現を網羅的に検索することも必要になる。

古谷, 寺井 (2018) は, REX-J コーパス (Spanger, Yasuhara, Iida, and Tokunaga 2012) に含まれる参照表現において比喩が用いられているかを検討した。須堯, 寺井 (2019) は, 青空文庫に出現するデータから「ように」「ような」を比喩指標要素として 19,209 文の自動抽出を試みた。抽出には係り受け解析器 (CaboCha) の出力に基づくパターンと分類語彙表番号に基づく喩辞の制約を課している。但し, 我々の研究のように比喩と例示の区別はされていないと考えられる。

また, 質問紙調査による指標比喩に対する小規模な評定調査が, 心理学の分野で進められている。中本, 楠見 (2004) は直喩形式の評定を行うために, 「A は B のようだ」に修正可能な表現を 500 以上収集した。収集元は, 心理学研究の先行研究 (楠見 1995; Nakamoto 2003; Ortony, Vondruska, Foss, and Jones 1985) や比喩表現辞典 (榛谷 1988; 中村 1995) であった。このうち 120 用例を取りあげ, 理解可能性・構成語類似性・独創性・面白さについて 7 件法により 60-64 人規模で評定調査を行った。平, 中本, 楠見 (2007) は直喩文 30 文に対して, 理解容易性・主題と喩辞の類似性・表現の意外性・表現の親しみやすさ・面白さについて 5 件法で 70 人規模で評定調査を行った。岡, 大島, 楠見 (2019) は, 中本, 楠見 (2004) の 120 用例について, 50 人規模で解釈産出課題を行い, 多義性の検討を行った。さらに 24 人規模で, 9 件法による表現の選好性課題・喩辞の慣習性課題 (Bowdle and Gentner 2005; Utsumi 2007) を行った。

**結合比喩関連:** 一方, 比喩であることが比喩指標要素によって明示されない比喩表現の収集は, 指標比喩と比べて困難である。このうち, 結合比喩は, 基底となる語義の体系を決定したうえで, 語義の接続・共起のちがいを手掛かりとした抽出が試みられている。なお, この語義の接続・共起のちがいは, 指標比喩においても有用な手掛かりであると考えられる。作業手順としては, 基本的に 2.1 節に示した, MIP・MIPVU に基づく手法が利用される。

日本語でも, MIP を用いた比喩表現コーパスの構築を目指し (伊藤 2014), MIP, MIPVU による比喩判定も試みられている (宮澤, 吉田, 宮尾 2016)。宮澤 他 (2016) は, MIP・MIPVU に基づき, 『岩波国語辞典第五版タグ付きコーパス 2004』を語義の判断基準とし, 目的語・動詞の結合についての比喩性判断 (中村の結合比喩に相当) を 1,100 文に対して行った。動詞の複合語の語義判断については, 『複合動詞レキシコン』 (神崎 2013) による。

これに対し、本研究は MIP, MIPVU を BCCWJ のコアデータおよび『分類語彙表』(国立国語研究所 2004) の語義に基づき整理を行う (3.1 節)。分類語彙表の語義情報である分類番号が付与されたデータ BCCWJ-WLSP (加藤, 浅原, 山崎 2019b) を基底となる語義の体系とし, 2.2 節に示す中村 (1977) に基づく比喩指標要素と語義の接続・共起に基づく比喩表現の収集作業を行った。さらに, 全文を対象として比喩指標要素を用いない, 語義の接続・共起のみに基づく, 結合比喩の収集作業についても着手している (Kato and Asahara 2019)。

その他: 紙の辞典形式の比喩表現用例集として, 榛谷 (1988), 中村 (1995, 2007), 野内 (1998, 2005), 佐藤, 佐々木, 松尾 (2006) などがある。

### 3 指標比喩データベースの構築手法

#### 3.1 MIP に基づく日本語指標比喩データベース整備作業の全体像

本節では指標比喩データベース整備作業の全体像について解説する。比喩用例収集手順として 2.1 節に示した MIP (Pragglejaz Group 2007), MIPVU (Steen et al. 2010) を採用し, 日本語対応した。本研究の比喩表現の捉え方は 2.2 節に示した 中村 (1977) に基づく。本研究の基本的な対象は, 比喩指標要素の近傍に, 非慣用, 非論理的など, 基本義の用法から逸脱した異例の結合を, 何らかの比喩的転換として確認できる用例である。しかし, 指標比喩の用例は概ね結合比喩と重複するが, 言語形式が認められても結合の確認が困難な例や, 言語形式 (指標と認定する範囲) が不明瞭である例, 結合要素そのものよりも周辺語句の影響が強くとめられる例などもあるため, 関連語を漏れないよう収集する必要がある。MIP を用いることで, 語単位に比喩性の判定が可能となり, 広く比喩表現に関係する要素の収集が可能となる。また, 比喩的な転換の見られる要素の結合に着目し, 慣用からの逸脱を確認する点においても, 基本義との対照を行う手法が有用であると考えられる。但し, この比喩的な転換の認識は, 主観的なものであり, 客観性を持たせることは本質的に困難である。

以下では, 図 3 に, 図 1 に示したオリジナルの MIP に対応する形式で, 日本語において言語資源を援用しながら客観性を担保して比喩表現を抽出する指標比喩データベース構築手法を示す。日本語比喩表現抽出手順の概要について示す。まず最初にテキストを読んで語義の理解を行う (1)。比喩用例収集にあたっては 中村 (1977) の 359 種類の比喩指標要素とその類似用例 (『分類語彙表』(国立国語研究所 2004) により判定) を手掛かりとし, その近傍を調査する。本研究の前提となる比喩指標要素の詳細を 3.2 節に示す。比喩用例収集対象を 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 文とする。比喩用例収集対象の一部には『分類語彙表』(国立国語研究所 2004) の語義 (分類番号) が付与されている。また, 同じサンプルに助動詞の用法が付与されている。本研究で語義の転換の認識に用いる語義・用法アノテーションの詳細を 3.3 節に示す。次に単位を決める (2)。本

- (1) 手掛かりとなる比喩指標要素の近傍にあるテキストを読んで文意の理解を行う（比喩指標要素の詳細については 3.2 節を参照）。
- (2) 分析単位を決める．文脈比喩以外の語義の分析単位として国語研短単位を基本とする。
- (3) 各単位の語義の認定を行う．語義の認定にあたっては『分類語彙表』の分類番号を基本とする。
  - (i) 基本的にテキストを読んで作業を行うが，3.3 節に示す語義アノテーション（加藤 他 2019b）が利用できる場合には語義アノテーションを援用する．語義の認定基準は（加藤 他 2019b）を参照されたい。
  - (ii) 基本義の同定にあたっては
    - 『日本国語大辞典』（concrete, precise, historical order）
    - 分類語彙表における代表義（山崎, 柏野 2017）
    - 語義アノテーション（加藤 他 2019b）における頻出義を参考にする。
  - (iii) 結合の比喩性の認識にあたっては，
    - 基本結合  
要素の基本義・代表義を認定したうえで，字義どおりに解釈できる要素の結合を指す。『動詞の意味・用法の記述的研究』（宮島 1972）に記述されている「意味的に動詞と結合する語の制限」に相当する。
    - 比喩的結合（中村 1977）  
要素の基本義・代表義を認定したうえで，字義通りに解釈できない要素の結合を指す。中村（1977）の B 型把握（言語単位の結びつきにおける慣用からの逸脱）に相当する。  
などを参考にする。
  - (iv) 文脈比喩の検討を同時に行う。
- (4) (i) 上記 (1)(2)(3)(i)～(iv) の手続きに基づき，比喩的な語義の転換・逸脱が認められた場合に「比喩」と認定する（3.4 節）。
  - (ii)
    - a. 「比喩」と認定した場合には，喩辞・被喩辞を同定し，その近傍の比喩的結合を抽出する。
    - b. 喩辞・被喩辞間もしくは結合の概念的写像である概念マッピングの類型や比喩種別を付与する。
    - c. 手掛かりとなった比喩指標要素についても記録する（3.5 節）。
  - (iii) また，非専門家の評定情報をクラウドソーシングにより収集する（3.6 節）。

### 図 3 MIP に基づく日本語比喩表現指標比喩データベース整備作業

作業では国語研短単位を基本とした。次に各単位の語義を推定する (3)(i)。基本的にはテキストを読んで作業を行うが，語義アノテーションが利用できる場合には語義アノテーションを援用する。さらに各単位について他の文脈での基本義との対照を検討する (3)(ii)。基本義の同定においては、『日本国語大辞典』<sup>7</sup>のほか，代表義（山崎, 柏野 2017）・語義アノテーションの頻出語（加藤 他 2019b）などを手掛かりとする。また要素の結合の比喩性の認識にあたっては基本結合（宮島 1972）・比喩的結合（中村 1977）を確認する (3)(iii)。上記 (1)(2)(3)(i)～(iv) の手続きに

<sup>7</sup> <https://japanknowledge.com/lib/search/nikkoku/>

基づき、比喩的な転換のみとめられる結合と判定し、比喩性を認定し、喩辞の出現箇所の同定を行う(4)(i)。なお、提喩など比喩的結合のみとめられないものも、その喩辞の出現箇所同定および比喩表現該当部の同定を行う。比喩表現には、比喩指標要素・比喩表現該当部・喩辞・被喩辞・結合・類型化・比喩種別(擬人・擬生や換喩・提喩・慣用など)を付与する(4)(ii)。詳細については3.5節で述べる。最後に、比喩表現に対する非専門家の評定値をクラウドソーシングにより収集する(4)(iii)。3.6節に評定値の収集手法について示す。

なお、MIPはMetaphor(隠喩、中村(1977)の結合比喩・文脈比喩)を主対象とした手法である。MIPを拡張したMIPVUは、比喩指標要素に該当する「MFlag」の付与も行っているが、指標比喩をMetaphorと区別するものではない。本作業は、指標のあるいわゆるSimile(直喩、中村(1977)の指標比喩)を主対象とするが、概ね同時に結合比喩が確認できる(比喩的な転換に差がない)ことから、収集した表現がMetaphorと別種とは位置づけない。また、類似に基づく典型的な転換ではない、いわゆる換喩(事物・事象の隣接性という類縁関係に基く「質的転換」)や提喩(類と種という概念の「量的転換」)のような用例も、要素の結合における比喩的転換の把握によって取得される<sup>8</sup>。このほか、比喩指標要素を用いることで、MIPにおいても作業効率の上がる可能性を期待した。いずれ比喩指標要素と共起しない第2類結合比喩・第3類文脈比喩も収集するが、本稿では網羅的な比喩データベース構築の出発点として、また指標比喩用例の実態分析を進めるために、3.2節に示す中村(1977)の比喩指標要素を手掛かりとした作業を行う。さらに比喩指標要素については、分類語彙表番号に基づく同義性を用いた汎化も試みる。

### 3.2 中村(1977)の比喩指標要素と比喩指標要素に対する分類語彙表番号付与

本研究では、受容過程における比喩性把握に着目し、比喩意識の指標となる言語形式を有する比喩(指標比喩)のデータベース構築を目標とする。「あたかも」「まるで」「よう」など、手掛かり句である「比喩指標要素」が明示的に出現する、指標比喩の事例収集を行うことになる。但し、指標は、比喩性を感じ取る(A型把握)という点において有用な言語形式であるが、比喩性の判断においては、指標比喩も結合比喩や文脈比喩と同様の手順を必要とする。よって、指標から当該表現に比喩性があるという可能性を感じるとしても、比喩性の判断には根拠を要する。そこで、本データベースにおいても、結合比喩・文脈比喩と同様、要素の結合や文脈における慣用からのずれに着目した情報を含める。そのために『分類語彙表』(国立国語研究所2004)の分類番号(分類語彙表番号)を語義の判断基準とした分析を進める。

我々は、指標比喩データベースの整備のため、中村(1977)の指標比喩の1,617類型と359種類の比喩指標要素を電子化し、分類語彙表番号を新たに付与した。紙面から手入力し電子化ファ

<sup>8</sup> 「換喩」「提喩」の定義は(中村1977)に基づく。

イルを構築し, 形態素解析 MeCab と UniDic を用いて, 短単位形態論情報を付与した. その後, UniDic 語彙素-分類語彙表番号対応表 (WLSP2UniDic)<sup>9</sup>(近藤, 田中 2020) を用い, 分類語彙表番号を付与した. 自動で付与できなかったものについては, 人手で修正を施した. 中村 (1977) の類型は分類語彙表番号順に整理されており, 本指標整備時から分類語彙表の体系を意識していたことが推察される.

比喩指標要素は, その品詞から D 類 (動詞)・F 類 (副詞)・J 類 (助詞)・K 類 (形容詞・形容動詞)・M 類 (名詞)・R 類 (連体詞)・S 類 (接尾辞) に類型化されている. 表 1 に例を示す. 「ようだ」以外にも多様な比喩指標要素が定義されていることがわかる. 比喩指標要素は, 類・種・号により階層化されている.

表 2 に比喩指標要素に対する分類語彙表番号割り当ての例として「よう (比喩指標要素の類型: K-9-1・UniDic 語彙素: 様)」を示す. 多義である「よう」は, 分類語彙表番号として 3.1300 の「様相」と 3.1130 の「類似」の 2 つを持つ. また, 同範囲には助動詞「ようだ」の用法として, 「類似」・「内容指示」・「例示」・「婉曲」の 4 つのいずれかが付与されている. このうち, 3.1130 「類似」の例のみが比喩表現に相当している. この分類語彙表番号を用いて, 多義語に対する語義の曖昧性解消を行いながらの比喩指標要素の抽出が可能になる. また分類語彙表番号が同じ要素を類似比喩指標要素とし, 候補の抽出に用いる.

表 1 比喩指標要素の類型と例

比喩指標要素の類型	例
D (動詞)	感じる, たとえる, 気がする, 似る
F (副詞)	まるで, いわば, 仮に, ではあるまいし
K (形容詞)	近い, 同じ, 同類, その通り, よう
R (連体詞)	へたな, 大した, いわゆる, 一種の
M (名詞)	もの, ようす, 役目, たぐい, 錯覚
J (助詞)	ほど, でも, さえ, というものは
S (接尾辞)	もの, 色, 様, ばり

表 2 比喩指標要素に対する分類語彙表番号割り当ての例: 「よう」分類番号 3.1300 (様相) と 3.1130 (類似)

類	部門	中項目	分類項目
相 (3)	抽象的關係 (.1)	様相 (.13)	様相・情勢 (.1300)
相 (3)	抽象的關係 (.1)	類 (.11)	異同・類似 (.1130)

<sup>9</sup> <https://github.com/masayu-a/wlsp2unidic>

### 3.3 『現代日本語書き言葉均衡コーパス』に対する分類語彙表番号アノテーションおよび助動詞用法アノテーション

比喩表現は意味的な転換を含む表現であるため、語義アノテーションを含むデータを用いることにより、効率的に抽出できる。また比喩指標要素も多義性があるため、助動詞の用法も含めた語義アノテーションを用いることにより、作業者の負担を減らすことができると期待される。本研究では語義アノテーションとして BCCWJ コアデータの一部に対する分類語彙表番号アノテーションデータ (加藤 他 2019b) を用いる。

同データは、前述の UniDic 語彙素-分類語彙表番号対応表 (WLSP2UniDic) により、BCCWJ の言語単位 (短単位・長単位) に対応可能性のある分類語彙表番号を列挙したうえで、人手で文脈上の正しい語義を選択したものである。列挙された分類語彙表番号の選択肢に該当する意味分類がない場合には、新たに適切な番号を付与してある。BCCWJ コアデータにはアノテーションの優先順位<sup>10</sup>が規定されており、レジスタ毎に A, B, C, D, E までの 2-5 の集合に分割されている。分類語彙表番号はサンプル PB (A), PB (B), PM (A), PM (B), PN (A), PN (B) (書籍・雑誌・新聞) 347,094 語 15,969 文の短単位すべてに人手によって付与されている。

表 3 に分類語彙表番号アノテーション例を示す。この例では、比喩指標要素である「よう」に、「関係-類-異同・類似」を示す「.1130」の分類語彙表番号下 4 ケタが付与されており、続く「闇を体験する」という結合から、比喩表現であるとされる。さらに喩辞「闇」の語義 1.5010 (体-自然-自然-光) が、被喩辞「サブリミナル」の語義 1.3001 (体-活動-心-感覚) に転換されて

表 3 BCCWJ に対する分類語彙表番号アノテーション

表層形	分類番号	分類
サブリミナル	1.3001	体-活動-心-感覚
の		
よう	3.1130	相-関係-類-異同・類似
に		
闇	1.5010	体-自然-自然-光
を		
体験	1.3050	体-活動-心-学習・習慣・記憶
し	2.3430	用-活動-行為-行為・活動
て		
いる	2.1200	用-関係-存在-存在
わけ	1.1113	体-関係-類-理由・目的・証拠
だ		

(サンプル ID : PB12\_0000, 『闇を歩く』中野純 (著))

<sup>10</sup> <https://github.com/masayu-a/BCCWJ-ANNOTATION-ORDER>

いることもわかる。

上に述べた分類語彙表番号は自立語が主な対象であるが、付属語の情報として、同範囲に助動詞の文法的な用法が別途付与されている(加藤, 浅原, 山崎 2019a)。以下の例文(4)では、2つ出現する「よう」のうち、前者は比喩指標要素になる「類似」用法であり、後者は「内容指示」用法である。助動詞の用法アノテーションにより、後者の比喩指標要素の語義の判定が不要になる。

- (4) 転機はひとみさんが六年生のとき。研究のため家族四人そろって渡米した。不登校は相変わらずだったが、本に興味を持ってくれた。日本から送った古典の現代語訳作品をはじめ、本棚に並ぶ本をむさぼり読んだ。一緒に行動しようと自分が好きな舞台に連れ出すと、長時間のオペラを食い入るように見つめていた。帰国後、中学校の校長は「特別扱いはできないが、卒業証書は出す」と言った。不登校に加え、何度となくリストカットも繰り返す。カウンセリングにも通った。本格的に小説を書くようになったのもこのころだ。

(サンプル ID : PN4g.00003<sup>11</sup>, 西日本新聞, 下線部は著者による)

本データは BCCWJ のコアデータの一部であるが、語義の曖昧性が解消されており、指標比喩データベースの整備作業の補助に利用できる。コーパスに分類語彙表番号と助動詞の用法が付与されているため、あらかじめ判明している指標となりえる要素(自立語もしくは助動詞)については、多義語が比喩指標要素に適合するか否かの絞り込みを行える。また、3.4 節で述べるように、分類語彙表の同じ分類番号を持つ語のグループを類義語としてみなし、中村(1977)の比喩指標要素の類義語を新しい手掛かり語として展開する。中村(1977)の比喩指標要素は分類語彙表番号順に整理されていることから、分類語彙表番号が抽出の手掛かりになると考える。さらに、自動処理により比喩を検出する場合にも、UniDic 語彙素-分類語彙表番号対応表(WLSP2UniDic)により、特徴量として分類語彙表の情報を用いることを想定する。

### 3.4 比喩表現出現箇所同定

次に比喩表現出現箇所同定の詳細について示す。

中村(1977)の比喩指標要素を BCCWJ コアデータ全体 1,290,060 語 57,256 文に対して語彙素のパターンマッチにより枚挙した。抽出された比喩指標要素の前後文脈 100 形態素を与え、言語学の知識を持つ者 2 名により比喩表現であるか否かの判定を行った。

さらに、分類語彙表番号で類義語句を展開したうえで、比喩表現候補を枚挙した。中村(1977)の比喩指標要素に付与された分類語彙表番号で、BCCWJ コアデータの分類語彙表番号付与済

<sup>11</sup> BCCWJ におけるサンプルの識別子。

表 4 比喩指標要素による候補数と抽出用例数

比喩指標要素	抽出用例数	候補数	抽出用例／候補
D (動詞)	323	33,231	1.0%
F (副詞)	113	33,070	0.3%
K (形容詞)	444	1,707	26.0%
R (連体詞)	0	1,840	0.0%
M (名詞) + J (助詞) + S (接尾辞)	84	5,729	1.5%
分類語彙表番号	56	21,541	0.3%

み部分 347,094 語 15,969 文 (書籍・雑誌・新聞) に対してパターンマッチを行い、前後 100 形態素を展開し、比喩表現であるか否かの判定を行った。例えば、D1 類については、「2.3001 (感覚)」「2.3061 (思考)」「2.3062 (注意・認知)」「2.1130 (類似)」「2.3066 (判断・推測)」「2.1310 (風・観・姿)」「2.3103 (表現)」などが対応する分類語彙表番号であった。この分類語彙表番号により展開し、短単位では 3,060 例、長単位では 539 例が候補として得られた。この分類語彙表番号に基づく用例では、短単位で 1.3% にあたる 41 例、長単位で 0.7% にあたる 4 例が比喩表現を含んでいた。

表 4 に、D 類 (動詞)・F 類 (副詞)・J 類 (助詞)・K 類 (形容詞・形容動詞)・M 類 (名詞)・R 類 (連体詞)・S 類 (接尾辞) および分類語彙表番号により展開された比喩指標要素によって展開された候補数と抽出用例数を示す。K (形容詞・形容動詞) は「よう」「みたい」などの率が高い要素が多い一方、R (連体詞) は 1 件もなかった。分類語彙表番号により汎化したパターンのうち、実際に比喩表現の指標であったものは 0.3% であった。

基本的に 1 つの比喩表現該当部に 1 つの喩辞が出現するように抽出する<sup>12</sup>。1 つの比喩用例が複数の比喩指標要素を含む場合 (複合用例) もあり、その場合は重複して確認するが、1 つの比喩表現該当部として抽出する。さらに、目的の比喩指標要素ではない要素による比喩表現該当部 (指標外比喩用例) も抽出した。このため比喩指標要素と比喩表現が必ずしも 1 対 1 対応するわけではない。結果、822 件の比喩表現該当部が抽出できた。このうち、複数 (2 つ以上) の比喩指標要素を含む複合用例は 157 件で、指標外比喩用例は 10 件であった。

### 3.5 比喩関連情報アノテーション

抽出された 822 件の比喩表現に対して、言語学の知識を持つ者 2 名により関連情報のアノテーションを行った。表 5 に比喩情報アノテーション例を示す。この事例を用いて比喩関連情報アノテーション作業の概要を説明する。

作業者は、3.4 節の作業により比喩表現出現箇所同定を行う際に、比喩指標要素の前文脈 100

<sup>12</sup> 但し、被喩辞・喩辞が明確でない提喩や文脈比喩が単体で出現する場合においてはこの限りではない。

表 5 比喩関連情報アノテーション例

前文脈	比喩指標要素	後文脈	
」とよばれ、うやまわれてきた。三原山の噴火が、観光客をよび、鳥を豊かにしてきたのだ。その三原山が、十五年前、五百年ぶりに大噴火をおこした。十カ所以上の火口から、溶岩がふきだしたのだ。流れ出た溶岩の跡が、幅百メートル、長さ三キロメートルもの大ききで、今も残っている。その姿は、	まるで	巨大な黒へびのようだ。ここを訪れた観光客は、当時のすさまじさにふれて、息をのむ。ゴツゴツとした黒い地面が広がり、ところどころにまっ黒にこげた木の切れはしが落ちている。火口だった場所は、すり鉢のように深くえぐれている。けれど、今はそこに緑の木がのび、時の流れを感じさせる。大噴火のとき、島民一万人全員を島の外に	
比喩表現該当部			
流れ出た溶岩の跡が、幅百メートル、長さ三キロメートルもの大ききで、今も残っている。その姿は、まるで巨大な黒へびのようだ。			
結合	被喩辞	喩辞	備考
跡が蛇	流れ出た溶岩の跡 (その姿)	巨大な黒へび	まるで・よう
類型 A	付属語 A	類型 B	付属語 B
跡	ハ	へび	ノヨウ
1.1720		1.5503	
種別	比喩指標要素分類		
擬生	F-1-1・K-9-1		

形態素と比喩指標と後文脈 100 形態素 (表 5 上部) の比喩用例を見て、何らかの要素の結合において慣用 (基本義) からの逸脱を含むと判断した喩辞 (喩える語) の位置を同定し、その文脈を比喩表現該当部として抽出する。提喩などで、喩辞が明確化されない場合は、提喩の出現位置を同定し、その文脈を比喩表現該当部として抽出する。この比喩表現該当部に対して、次に示す結合・被喩辞 (喩えられる語)・喩辞・類型・種別・比喩指標要素分類などの情報を付与する。

**結合:** 結合は、被喩辞相当句が文脈上語義的に逸脱しない文脈において、喩辞相当句が語義的に逸脱して位置する場合、比喩的結合として逸脱が確認できる語句の接続を抽出する。用例によっては、複数の結合を含む。要素の結合の欄には「跡が蛇」という結合に比喩的な語義の転換があると判定した旨を記載する。但し、例示や指示に近い用法 (境界事例)、指標がなければ成立しない表現などは、可能な形で記述し、種別欄にその旨を付記する。また、提喩のように本質的に比喩的結合がみとめられない事例<sup>13</sup>もある。基本結合か否かの判定には、分類語彙表に基づく動詞の結合価の大規模な資料である 宮島 (1972) を参照する。また、中村 (1977) の 305 ページ以下に比喩的結合 5,537 例<sup>14</sup>が示されており、これらも参考にする。

**被喩辞・喩辞:** 被喩辞は「喩えられる語」、喩辞は「喩える語」を意味する。表 5 では、被喩

<sup>13</sup> 例えば、「空気のような存在」「どこかこの世ではない場所」のような提喩の例などでは、用例中に被喩辞・喩辞が明示されないため、比喩的結合も確認できない。

<sup>14</sup> 本作業に先立ち、中村の比喩的結合についても MS Excel 形式に入力したうえで、分類語彙表番号を付与した。

辞である「流れ出た溶岩の跡（その姿）」<sup>15</sup>と喩辞である「巨大な黒へび」を抽出する。比喩表現（概念マッピング）であることを読み取るにあたり、喩辞や被喩辞は結合から直接取得できない（「時代が痩せる」など）例が多いが、喩えるものと喩えられるものの認識が必要である。要素の結合に何らかの比喩的な語義の転換があることを明確化するとともに、概念マッピングの類型化の補助として収集する。比喩用例中に明記がなければ前後文脈から取得するか文脈を整理して記入する<sup>16</sup>が、先に述べた提喩をはじめ、いずれも不明瞭で取得できない場合がある。

**類型：**類型は、比喩的結合に対して、字義どおりの解釈ができない（つまり、基本結合から逸脱している）ことを明らかにするための形式である。結合を要素に分解したもので、結合に含まれる喩辞と接続する語との間には比喩となる語義的な転換があり、その語義的な転換が分類語彙表上のどの語義とどの語義が基本結合から逸脱して接続しているかを示すものである。喩辞の語義を被喩辞の語義に置き換えると基本結合になることを想定する。類型には、語義の転換のある結合に相当する要素とその分類語彙表番号、付随する付属語の情報（結合の関係）を記載する。なお、結合要素が固有名詞をはじめ複雑な構成を有した表現であるなどの場合には、分類語彙表番号が付与可能なレベルに概念化する。この類型の分類語彙表番号の差異から、どのような語義の間に転換が起きているのかが観察できる。さらに、電子化した中村の比喩的結合 5,537 例と対照することも可能である。

**種別：**種別には、擬人化・擬物化・擬生化・具象化など、比喩性の判断根拠となった要素の結合分類、喩辞と被喩辞の関係にあたる概念マッピング情報を付与する。

- a). 『擬人』化（物（抽象物・具体物）を人に喩える）  
例「溶岩が島民をあざ笑う（かのように）」「時代が痩せている（みたい）」
- b). 『擬物』化（人を物に喩える）  
例「私を貴方の道具としてお使いください」「平七郎が彫像のように立つ」
- c). 『擬生』化（動植物に喩える、活喩、準擬人化）  
例「生き物としての音楽」
- d). 上記外の『具象』化（抽象物を具体物に喩える）  
例「学術研究の基盤（として）」「心を洗われた（と感じる）」
- e). 上記外の『抽象』化（具体物を抽象物に喩える）  
例「何かの啓示のように雲が浮かぶ」
- f). その他の『転換』（上記外の別種の物事に喩える）  
例「被害想定は防災カルテ（として）」「マエストロを歌舞伎役者にたとえた」

<sup>15</sup> ここでは、テキスト中において同定可能な語句として抽出している。なお、MIP の手順により、複数の作業者が各々比喩性を感じた短単位全てへのマークを別途行っている。

<sup>16</sup> ゼロ代名詞や省略など、比喩表現該当部に被喩辞が文字列として出現しない場合は、「#」を付記し、喩えられているものを記述する。

上記 a)~e) は、類似性に基づく転換の、人(生物)・物体・抽象体における大まかな概念マッピングであり、中村(1977)の結合類型の9割以上が上記 a)~e) に分類される。しかし、他の事物・事象に置き換える際には、それほど大きく移行しない場合も見られる。たとえば、液体を気体に喩える(体の転換)、聴覚を視覚に喩える(感覚の転換)、哺乳類を鳥類に喩える(界の転換)などの転換が含まれる。本作業においては、このような小規模な転換の出現頻度の少ないことが予想されたため、その他の「転換」としてまとめることにした。但し、後に付与する分類語彙表番号の中分類などにより、詳細な分類が可能である。なお、MIPVU では personalization (擬人化) を定義する<sup>17</sup>が、日本語の比喩の議論で言及される、擬物化・具象化・抽象化といったものは導入されていない。

以上の一般的な類似性に基づく転換のほか、語感のずれなどで感じやすい類と種の関係(量的転換)を判断根拠とした場合は『提喩』と分類し、付属物や中身を用いてあるもの全体を表現しているなど、隣接性(質的転換)を判断の根拠とした場合には『換喩』と分類した<sup>18</sup>。文脈比喩である(結合要素からのみでは判定できない)場合は『文脈』に分類し、慣用表現と判断された場合などは『慣用』に分類した。指標要素が別用法である可能性も考えられた場合は、「例示」「指示」「非断定」などの用法を『その他』として付記した。

これらの種別情報は、被喩辞-喩辞間・結合単位<sup>19</sup>に付与を行うため、複数の結合を持つ1つの比喩表現に対して複数の種別情報を付与する場合がある。例えば、「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...」では、結合「イカが素麺」(『転換』)と結合「素麺のように細い」(『提喩』)の2つの結合に対して付与し用例全体には『転換-提喩』とタグ付けする<sup>20</sup>。また、「肝臓はすっかりブルーになっていたのか」という表現においては、「肝臓(の状態)」に対して『提喩』を付与するほか、結合「ブルーにな(る)」(『換喩』)と結合「肝臓がブルーにな(る)」(『擬人』)の2つの結合に付与し、用例全体には『提喩-換喩-擬人』を付与する。さらに1つの結合が複数の種別ラベルを持つ場合がある。例えば、「その事態は会社がつぶれるようなもので」という表現中の結合「会社がつぶれる」に対しては、『具象-換喩-慣用』の複数ラベルが付与される。

**比喩指標要素分類・備考:**最後に、出現した中村(1977)の比喩指標要素の分類記号「F-1-1(副詞:まるで)・K-9-1(形容詞:よう)」を付与する。比喩性に関わると考えられた指標要素全てにタグ付けする。備考には関連する情報を付与する。喩辞と被喩辞を含め、比喩性に関わると考えられた語句を可能な限り収集した。

<sup>17</sup> MIPVU の定義では擬人化に擬生化の一部が含まれている可能性がある。

<sup>18</sup> 『換喩』・『提喩』(『メトニミー』・『シネドキシ』)は研究者により、指すものが異なる傾向がある。本研究における分類は中村(1977)に基づく。

<sup>19</sup> 『提喩』などは単体の表現に、『文脈』はより広い範囲の表現に付与することもある。

<sup>20</sup> 比喩表現該当部全体に対して付与する場合には、結合(要素)間の種別の共起を“-”で、結合(要素)内の種別の共起を“\_”で表す。

### 3.6 クラウドソーシングによる評価

これまでの作業は言語学の背景知識を持つ者により進めたが、専門的知識を持たない者が指標比喩表現をどのように捉えるかは明らかでない。そこでクラウドソーシングにより、表現に関する3つの観点（比喩性・新奇性・わかりやすさ）と、3つの観点との関連性を調査するための2種（擬人化・具体化）を加えた5つの評価を行った。なお、表現に関する同様の質問紙調査を中村(1977)の例文（比喩表現のみ）やIPAL辞書の動詞・形容詞の例文（比喩表現以外も含む）にも行った。

- 比喩性：何かを他の物事でたとえる（比喩）表現を含むか<sup>21</sup>  
提示する例文はすべて比喩表現を含んでいるが、実験協力者が比喩であることを認識できるかについて評価する  
質問項目：「何かを他の物事でたとえる（比喩）表現をふくんでいますか。」
- 新奇性：新しい表現を使っていると思うか  
（比喩）表現の新しさについて評価する。  
質問項目：「新しい表現を使っていると思いますか。」
- わかりやすさ：わかりやすく表現されているか  
（比喩）表現を導入することによって、文意がわかりやすくなっているかわかりにくくなっているかを評価する。  
質問項目：「わかりやすく表現されていると思いますか。」

以下は、比喩表現に多いとされる「擬人化」「具象化（具体化）」が一般に読み取られるのかを調査するために用いた観点である。これらの概念マッピングを読み取ることで、その他の観点の判定に影響が現れるのかを調査するという補助的な目的で設定している。

- 擬人化：人でないものを人に見たてているか  
比喩であるとして、擬人化か否かが認識できるかを評価する。  
質問項目：「人でないものを人に見たてていますか。」
- 具体化：具体的なものに見たてて説明しているか  
比喩であるとして、具象化か否かが認識できるかを評価する。具象化が一般的な用語でないために「具体的」という用語を用いた。  
質問項目：「具体的なものに見立てて説明していますか。」

本調査はYahoo!クラウドソーシングを用いて行った。対象は20歳以上のYahoo!クラウドソーシングのアカウント所持者であった。調査協力者は、上記5つの観点について、「0：まったく違う」～「5：そう思う」の6種類の評定値（6件法）から1つ選択する。前節で抽出され

<sup>21</sup> 我々は比喩性を二値的なもの（比喩であるか否かの二律背反なもの）ではなく、程度性を持つものとして扱う。

た822件の比喩表現該当部を820件に整理したうえで提示文とした<sup>22</sup>。2019年2月15日8:00-2月16日00:10の間に異なり1,657人の評定値を収集した。5つの観点の組み合わせについて全く同じパターンの回答を5回以上行った実験協力者を、データから除外し、1,164人分のデータを分析対象とした。実験協力者は6用例を判定するごとに10円相当のTポイントを得、最大30用例まで評価した。費用は149,040円であった。

結果、1つの例文に対する22人~77人分の評定値（平均33.0人）を得た。

## 4 比喩指標要素近傍における比喩表現使用傾向の検証

### 4.1 アノテーションデータ

以下ではアノテーションデータの分布について確認する。

表6にレジスタ毎の比喩表現該当部の分布を示す。レジスタの文数・形態素数が異なるために10,000文あたりの相対度数、1,000,000形態素あたりの相対度数で評価する。書籍(PB)が最も多く、Yahoo!知恵袋(OC)や白書(OW)が少ない傾向にあった。

表7にレジスタ毎の種別の分布を示す。レジスタ毎に各種別の度数と相対度数(10,000文あたりの種別(結合)数を計数した。なお、「結合0」は「その他」に分類し、1件と数えた。種別は1つの事例に複数の結合を認定したうえで結合ごとに複数のラベルを付与する(マルチラベル)ため、レジスタ毎の総計は表6の比喩表現該当部数と異なる。種別を分類すると、「転換」(小規模な転換)が多数を占める(10,000文あたり57.8件)ことがわかる。文学作品の結合類型

表6 比喩表現該当部および結合の分布

	Y!知恵袋 (OC)	白書 (OW)	Y!ブログ (OY)	書籍 (PB)	雑誌 (PM)	新聞 (PN)	全体
比喩表現該当部数	25	48	63	299	198	189	822
(10,000文あたり)	40.9	82.4	89.2	308.9	157.9	117.8	143.6
(1,000,000形態素あたり)	226.7	210.4	537.4	1275.6	826.9	524.2	637.2
結合0の比喩表現該当部数	0	0	1	2	1	0	4
結合1の比喩表現該当部数	24	43	53	238	165	152	675
結合2の比喩表現該当部数	1	4	8	52	31	31	127
結合3の比喩表現該当部数	0	1	1	7	1	6	16
形態素数	110,280	228,172	117,242	234,400	239,440	360,526	1,290,060
文数	6,110	5,825	7,059	9,678	12,542	16,042	57,256

<sup>22</sup> 該当部だけでは比喩表現と認識されない場合は「<中略>」などとして整理を行った。結果複数の比喩表現該当部が1つに集約された場合があった。複数の比喩指標要素を持つ比喩表現該当部は複数回提示したが、同一の比喩表現該当部が同じ調査協力者に割り当てられないように配慮した。

表 7 結合に割り当てられる種別の分布

	擬人	擬生	擬物	具象	抽象	転換	換喩	提喩	文脈	慣用	その他
Y!知恵袋 (OC)	0	2	0	2	0	6	1	1	2	11	3
(相対)	0.0	3.3	0.0	3.3	0.0	9.8	1.6	1.6	3.3	18.0	4.9
白書 (OW)	1	0	1	24	0	5	1	1	0	23	0
(相対)	1.7	0.0	1.7	41.2	0.0	8.6	1.7	1.7	0.0	39.5	0.0
Y!プロゲ (OY)	1	4	4	16	1	22	5	13	1	11	1
(相対)	1.4	5.7	5.7	22.7	1.4	31.2	7.1	18.4	1.4	15.6	1.4
書籍 (PB)	<u>31</u>	20	20	66	0	147	16	45	2	32	19
(相対)	<u>32.0</u>	20.7	20.7	68.2	0.0	151.9	16.5	46.5	2.1	33.1	19.6
雑誌 (PM)	9	13	8	26	2	85	6	35	2	37	25
(相対)	7.2	10.4	6.4	20.7	1.6	67.8	4.8	27.9	1.6	29.5	19.9
新聞 (PN)	10	7	10	70	0	66	11	21	2	45	20
(相対)	6.2	4.4	6.2	43.6	0.0	41.1	6.9	13.1	1.2	28.1	12.5
総計	52	46	43	204	3	331	40	116	9	159	68
(相対)	9.1	8.0	7.5	35.6	0.5	57.8	7.0	20.3	1.6	27.8	11.9

注：(相対)行は、10,000 文あたりの種別（結合）数。

(中村 1977) では 1 割未満の種別であったことから、隠喩として一般的ではない転換であるとも考えられる。大きな転換であれば、カテゴリーの違いによって比喩性のあることが読み取りやすいが、小規模な転換であるがゆえに、指標を用いて比喩であるという把握を促す傾向が考えられる。比喩表現の使用実態における指標の必要性が考えられる。本データベースで収集した用例の分析によって、指標を用いる必要のある表現の整理や、例示や内容指示など比喩との区別が問題視される表現の傾向も明らかになると期待される。

擬人化は文学作品を含む書籍 (PB) に多く (表 7 下線部)、中村 (1977) 文学作品の結合類型の分布 (擬人化が全類型の 41.8%) の結果と整合する。今後、書籍サンプルに付与されている NDC (日本十進分類法) の情報と対照することで、書籍のジャンル毎の検討を行いたい。

また、抽象の用例が 3 件確認された。MIP, MIPVU のアノテーションの定義においては、具象化・具体化の方向で語義の転換があることを前提に比喩表現の認定がなされていたが、本データベースでは抽象化の方向でも語義の転換があることが認められた。今後、他言語においても抽象化の用例があるかを検討する必要がある。

表 8 に比喩指標要素の頻度上位 10 件とその種別の分布を示す。よく直喩指標で用いられると考えられている K-9-1 「よう」は、357 例<sup>23</sup> (今回抽出した指標比喩用例の 43%) に出現した。D-1-16 「する」は「にする」「とす」が含まれ、特に慣用表現が多い傾向にある。K-9-3 「みたい」と F-1-1 「まるで」は小規模な転換 (『転換』) が多い一方、D-1-1 「感じる」は『具象』が多

<sup>23</sup> 種別は 1 事例に複数付与されるために行の合計より少ない数になる。

表 8 比喩指標要素 (頻度上位 10 位まで) と種別の分布

指標記号	指標要素	度数	擬人	擬生	擬物	具象	抽象	転換	換喩	提喩	文脈	慣用	その他
K-9-1	よう	357	36	27	19	61	3	183	12	53	1	14	46
D-1-16	する	223	7	3	10	79	0	17	11	9	1	124	2
K-9-3	みたい	61	1	5	3	3	0	40	2	15	1	0	4
F-1-1	まるで	47	3	5	2	3	1	29	1	4	1	0	4
K-11-1	である	38	3	4	4	2	0	21	0	6	0	2	0
M-1-2	もの	35	3	2	1	7	0	16	0	8	0	0	7
D-1-1	感じる	20	0	2	0	10	0	4	3	1	2	3	0
D-1-2	思う	18	1	0	0	4	0	8	1	2	1	4	2
S-1-5	…状	18	0	0	0	1	0	14	0	2	0	0	1
D-1-12	たとえる	14	1	0	2	2	0	7	1	1	0	0	1

表 9 種別ラベルの共起

	擬生	擬物	具象	抽象	転換	換喩	提喩	文脈	慣用	その他	<単独>
擬人	5	1	9	0	13	0	4	0	0	0	25
擬生		1	8	0	9	0	5	0	3	1	17
擬物			4	0	5	0	2	0	1	0	25
具象				0	23	5	14	1	24	8	90
抽象					0	0	1	0	0	0	2
転換						8	20	1	4	0	241
換喩							7	0	3	1	18
提喩								1	3	3	59
文脈									0	0	6
慣用										0	119
その他											56

かった。小規模な転換においては、「まるで～のような」などの指標を組合せた類型が用いられやすく、言語形式によって比喩性の把握を促す傾向が強いといえる。「感じる」は身体性に関わる比喩表現に用いられ(菊地, 加藤, 浅原 2018), 被喩辞が身体で感じることの可能な具体物にされることになる。

表 9 に種別ラベルの共起について示す。具象と転換 (用例 (5)), 具象と慣用 (用例 (6)(7)), 転換と提喩 (8)(9) が共起する傾向が見られた。具象化は一般に慣用化している例が多い。用例 (5) は「作品が文化に筋を通す」(具象) と「作品のさわやかな筋」(転換\_具象) の 2 つの結合がある<sup>24</sup>。用例 (6) は「胸が空く」(具象) と「胸のすくような成功」(慣用) の 2 つの結合がある。用例 (7) は「チャンスを手にする」(具象\_慣用) の 1 つの結合がある。用例 (8) は「得点で王となる」(転換) と「日本語キング (日本語クイーン)」(提喩) の結合がある。用例 (9) は「相

<sup>24</sup> 用例単位のラベルとして結合 (要素) 間の種別の共起を “-” で、結合 (要素) 内の種別の共起を “\_” で表す (再掲)。

撲取りみたいな体格」(提喩\_転換)の1つの結合がある。なお、3つ以上ラベルが共起する場合(用例(10):「政権が崩壊」(具象\_換喩)・「崩壊が目の当たり」(具象)・「崩壊を目の当たりにする」(慣用))もある。

- (5) 佐藤さんが指摘する「意地」は、大衆文化の中に、1本のさわやかな筋を通していているように思える。

(「具象-具象\_転換」サンプルID:PN4a\_00010)

- (6) 胸のすくような成功の軌跡を一挙公開

(「具象-慣用」サンプルID:PM31\_00275)

- (7) タイ・バンコクではチャンスも手にした。

(「具象\_慣用」サンプルID:PN4g\_00002)

- (8) 次回の試験は十一月です。最高得点者は日本語キング、日本語クイーンとして表彰されます。

(「転換-提喩」サンプルID:PN3b\_00009)

- (9) 兄は独り者で、相撲取りみたいな体格だという話だった。(「転換\_提喩」サンプルID:PM31\_00020)

- (10) 「私たちはイラクの中央政権の崩壊を目の当たりにしている」

(「具象\_換喩-具象-慣用」サンプルID:PN3a\_00002)

## 4.2 クラウドソーシングによる評価データ

### 4.2.1 クラウドソーシングによる評価値の基礎統計と上位事例・下位事例

表10にクラウドソーシングによる評価のレジスタに基づく提示用例ごとの平均評価値の平均(用例単位マクロ平均)を示す。比喩性は、書籍(PB)・Yahoo!ブログ(OY)・雑誌(PM)が高く、白書(OW)が低い傾向が見られた。新奇性は、Yahoo!ブログ(OY)・書籍(PB)が高く、白書(OW)が低い傾向が見られた。わかりやすさは、Yahoo!知恵袋(OC)が高く、白書(OW)が

表10 クラウドソーシングによる評価の分布(レジスタに基づく用例単位マクロ平均)

	OC 知恵袋	OW 白書	OY ブログ	PB 書籍	PM 雑誌	PN 新聞	総計
比喩性	2.30	1.66	2.80	2.93	2.76	2.50	2.69
新奇性	1.70	1.49	2.08	2.06	2.01	1.85	1.96
わかりやすさ	3.19	2.87	2.96	3.06	3.03	3.12	3.05
擬人化	1.43	1.29	1.54	1.56	1.45	1.48	1.50
具体化	2.61	2.41	2.77	2.93	2.88	2.79	2.83

低い傾向が見られた。擬人化は、書籍 (PB)・Yahoo!ブログ (OY) が高く、白書 (OW) が低い傾向が見られた。具体化は、書籍 (PB)・雑誌 (PM) が高く、白書 (OW) が低い傾向が見られた。これらの傾向に対しては、実験協力者のバイアスを考慮したうえで統計的に分析する必要がある。4.2.2 節で詳細に検討する。

表 11 にクラウドソーシングによる評定の種別に基づく提示用例単位マクロ平均を示す。比喩性は、擬生が最も高く、慣用が最も低い傾向が見られた。新奇性は、抽象が最も高く、慣用が最も低い傾向が見られた。わかりやすさは、転換が最も高く、抽象が最も低い傾向が見られた。擬人化は、擬人・擬生が高く、慣用が低い傾向が見られた。具体化は、擬生が高く、慣用が低い傾向が見られた。また、具象が低く、抽象が高いという傾向も見られた。これらの傾向に対して、実験協力者のバイアスを考慮したうえで統計的に分析する必要がある。4.2.2 節で詳細に検討する。

以下では、クラウドソーシングによる評定の上位事例・下位事例を検討する。

表 12 に比喩性の評定が上位の事例を示す。下線部は比喩指標要素を表す。比喩指標要素の範囲が重なる場合は最大の範囲のみ示す。なお、クラウドソーシングの呈示画面には、比喩指標要素が比喩性判断を促進しないように比喩指標要素を呈示しない。「顔／りんご」「海に浮かぶ様／一本の棒」「本／宝石箱」といった結合は論理的に成立しないため、非比喩の解釈が発生しにくいと考えられる。また、「赤さ」「線状」「高い価値」のような、喩辞の有する属性を根拠とする見立てであることが、比喩性の判断に影響した可能性がある。

表 11 クラウドソーシングによる評定の分布 (種別に基づく提示用例単位マクロ平均)

	擬人	擬生	擬物	具象	抽象	転換	換喩	提喩	文脈	慣用	その他
比喩性	3.19	3.49	3.23	2.37	3.20	3.11	2.27	2.99	2.41	1.69	2.74
新奇性	2.34	2.31	2.23	1.92	2.43	2.06	1.91	2.05	2.01	1.52	1.94
わかりやすさ	3.04	3.07	3.10	2.97	2.62	3.14	2.91	3.04	2.77	3.05	3.05
擬人化	2.24	2.22	1.94	1.38	1.47	1.50	1.45	1.46	1.66	1.23	1.44
具体化	3.00	3.17	3.06	2.53	3.05	3.10	2.51	2.98	2.55	2.25	3.00

表 12 比喩性の評定 (上位事例)

サンプル ID	比喩事例	比喩性
PB29_00013	顔は、りんごの <u>ように赤く</u> 、目は、大きく、まんまるで、頭でっかちな子どもです。	4.53
PB39_00023	泳ぎつかれたら一本の <u>棒みたいに</u> 海水に浮かび、	4.47
PN2a_00017	私はすぐ先の納屋に向かって蹣跚と足を運んだ。人が <u>見たら糸操りの人形のような歩きかた</u> だったろう。	4.43
PB40_00035	してみるとこの本は、間違いなく私の「 <u>宝石箱</u> 」なのである。	4.43
PM11_00260	化粧が <u>天才的に</u> ヘタでピカソ <u>みたいな</u> メーカー <u>になってる</u> 女や、	4.41

表 13 に比喩性の評定が下位の事例を示す。「口にする」「耳にする」「柱とする」「基礎にする」が比喩性評定の下位に来た。前二者は、発話や聴取をその行為が行われる器官を用いて表現する、いわゆる換喩である。一般に、換喩は「赤ずきん」「鍋を食べる」などの典型表現をはじめ、意識的に用いられる例が少ないと考えられる。また、「柱とする」「基礎にする」は (Lakoff and Johnson 1980) でいうところの ‘THEORIES ARE BUILDINGS’ の概念メタファーに基づく比喩表現である。人間の認知に根ざした比喩である場合、表現に対する異質性の印象が薄い可能性がある。なお、PB35\_00013 の例において、比喩指標要素「となり」と「という」は、比喩表現を構成する指標ではない。このような比喩指標要素は比喩性判断に影響を与えることが考えられるため、クラウドソーシングの際には呈示していない。

表 14 に新奇性の評定が上位の事例を示す。「home / 犠牲」「時代 / 痩せる」「リアリティーの無さ / スクリーンの向こう側」「闇 / サプリミナル」「人生 / ステップを刻みながら斜面を登る」といった取り合わせの異常性が目立つ。また、瞬きによる一瞬の闇を「サプリミナル」として捉え、死に至るまでの人生の歩みを「ステップを刻む」ものとして捉えるといった、通常なら

表 13 比喩性の評定（下位事例）

サンプル ID	比喩事例	比喩性
PB29_00003	三宝が佐倉の名を口にした時こそそのチャンスだったのだが、なぜか言いそびれたのである。	0.50
PB29_00026	アルコールをほとんど口にしなくなってずいぶんになる。	0.52
PB35_00013	「ここまでが標準仕様で、あとはオプションとなります」という言葉をよく耳にしませんか？	0.55
OW6X_00073	タイ及びその周辺国への薬物分析技術の移転等を柱とする「薬物対策地域協力プロジェクト」に、	0.56
PN4d_00004	「具体的な議論の基礎にするために提案した」	0.65

表 14 新奇性の評定（上位事例）

サンプル ID	比喩事例	新奇性
PB25_00063	近年この国では肝心の home がなおざりにされて、容器にすぎない house を追い求める傾向が強い。主客転倒して house を獲得するために home が犠牲にされることだって珍しくないのだ。	3.61
PM11_00002	なんか時代が痩せてるって感じがするよね。	3.45
PN1c_00007	クローン人間という言葉は未だにリアリティーを獲得できずにいる。まるでスクリーンの向こうのヴァーチャルな世界のこのように。	3.43
PB12_00001	そう考えると、私たちは瞬きするたびに、ほんの一瞬ずつ、サプリミナルのように闇を体験しているわけだ	3.33
PB25_00063	人は最終的な home (墓) にたどり着くまで、まるでステップを刻みながら (中略) 斜面を登るように営々と自分の居場所を探しつづけるしかないのだ	3.30

ば予測できないそれぞれの物事の側面を表現している。これが新奇性の一つの根拠になると推測される。

表 15 に新奇性の評定が下位の事例を示す。「手にする」「後にする」「教訓とする」「目の当たりにする」といった「Xに(と)する」形式と「動きを止められたように」が評定の下位となっている。「Xにする」形式は換喩的に意味を形成するものであり、積極的に表現対象となる事物をカテゴリーの異なる事物に見立てるものではない。「Xに(と)する」形式は比喩性の評定でも下位であったが、新奇性も薄い。「動きを止められたように」は助動詞「よう」の様態用法と区別しがたい境界的な例であり、比喩の新奇性という面では低いと判断されると推測される。

表 16 にわかりやすさの評定が上位の事例を示す。大人として求められる性格・性質を欠いた人間を「子供」と表現することは慣用的であるが、〈幼稚〉という明確な属性を付加することができる。このように明確に属性が用いられる場合、理解されやすいと推測される。また、「噴火／噴水」や「大きな鉄船／巨大な壁」の取り合わせは被喩辞の様態を「噴水」や「巨大な壁」といった視覚的なイメージを有する喩辞によって説明している。このように明確な視覚的イメージを用いた場合も理解しやすい比喩として判断されると推測される。「いつも黒い服を着ている女性」を「葬式帰りのように見え」と表現することは、「黒い服」が葬儀に着ていくことから、連想が容易であることによりわかりやすくなるのだと推測される。

表 17 にわかりやすさの評定が下位の事例を示す。「強たる物欲しみして身を亡すに譬給えるにや」は比喩関係の難しさよりも古語であるという文体による解読の難しさが存在する。また、

表 15 新奇性の評定（下位事例）

サンプル ID	比喩事例	新奇性
PN5a.00005	1 回の公演で手にするお金は、交通費約 2 千円。	0.64
PN3b.00014	衆院が解散し、国会を後にする小池百合子氏	0.67
OW6X.00040	すべての都道府県で、阪神・淡路大震災を教訓とした見直しを行っているが、	0.67
PB39.00013	兵はもとより庶民たちも動きを止められたように固まっていた。	0.70
PN23.00006	事件を目の当たりにした子どもたちは、今なお深い心の傷に苦しんでいる。	0.76

表 16 わかりやすさの評定（上位事例）

サンプル ID	比喩事例	わかりやすさ
PB39.00023	あの人子供みたいだね。	4.31
PB1n.00024	溶岩は、粘りが少ないため、噴火すると噴水のようにふきあがる。	4.28
OC09.00369	同じ部署にいつも黒い服を着ている女性がいて、全身黒で日傘やバッグも黒なので、葬式帰りのように見えます。	4.26
OW6X.00073	武器庫として使われたアパート	4.25
PB39.00009	大きな鉄船は巨大な壁のように目の前をふさいでいる。	4.20

「くずC」<sup>25</sup>や「ヒョーヒボハン（表皮母斑）」のように被喩辞が馴染みの薄いものであり、何が喩えられているのか分からない場合もわかりやすさが低くなると推測される。これらのわかりにくさは比喩の問題ではなく文体・指示物の問題である。「(処女の失い方が) 天使のような捨て方」であるという比喩は被喩辞である「処女の失い方」が明記されていないので、「天使」と「捨て方」の結合から理解することになる。処女についての文脈が無い場合、両者の結びつきは意外性があり、理解が難しくなる。「音楽としての意味思想の世界」は「音楽」に「意味思想の世界」が存在すると捉えている。ただ、喩辞・被喩辞の両者ともに抽象的であり、具体性が伴わないためにわかりやすさは低くなると推測される。

表 18 に擬人化の評定が上位の事例を示す。いずれの比喩表現も「月／王さま」「鳥／恋人」「水／友だち」「船／片眼・横たえる」「道路・公園／養子」「ボランティア／里親」と喩辞が人物・身体部位・動作であり、語彙的に人間であることが保証されている。人に喩えていることが語彙的に明確であることが擬人的であるという判断につながるのだと推測される。

表 19 に擬人化の評定が下位の事例を示す。擬人化的な比喩関係を持たない例が評定の下位となっている。また、「火口／すり鉢」のように場所の特徴を具体的な物によって喩えている例は

表 17 わかりやすさの評定（下位事例）

サンプル ID	比喩事例	わかりやすさ
OY14.02403	「此御歌は人の強たる物ほしみて身を亡すに譬給へるにや」	1.02
PB32.00008	また反対に、よし天使のような捨てかたであっても、それが甲でもよし乙でもよし、あるいは丙でもよかったというような場合を想像し得られませんか。	1.05
OY01.00848	そんな気がして、くずCをちょい宝くじ的に捨てて期先で↓を見るつもりなんですが？	1.16
PB54.00015	オビジョーニコマカナカッションクノイボガデキルヒョーヒボハン	1.27
OY14.20592	音楽が、言葉の介在しない音楽としての意味思想の世界を形成しつつ、確固たる意志のもとに紡ぎ合わさる。	1.29

表 18 擬人化の評定（上位事例）

サンプル ID	比喩事例	擬人化
PN2g.00004	月は大きくて王さまのようです。	4.21
PN1d.00001	最初は離れてとまっていたそうですが、やがて二羽が並び、写真のように恋人同士を思わせるしぐさに移ったそうです。	3.80
PB39.00009	この町の子どもたちはまるで魚のように水と友だちだった。	3.67
PB53.00056	荒海での厳しい操業を物語るかのように、片眼を失った船も一隻、しょんぼりと岸壁上に体を横たえていた。	3.56
PN2a.00016	道路や公園を「養子」に、ボランティアを「里親」にみたてて清掃美化活動をする。	3.56

<sup>25</sup> 分類語彙表番号 1.3721 資本・金銭：「C」は前場（前引け）や後場（大引け）の時刻にできた売買の値。

人であることが関係しないため下位になると推測される。

表 20 に具体化の評定が上位の事例を示す。「メーク／ピカソ (の絵)」「噴火／噴水」「海に浮かぶ様／一本の棒」「鉄船／巨大な壁」「内出血の赤さ／ペンキ」はいずれも喩辞が明確な視覚的イメージを有している。そのようなイメージを被喩辞と重ね合わせるにより、被喩辞の具体的な様態が受け手に伝わると推測される。「メーク／ピカソ」「海に浮かぶ様／一本の棒」は比喩性判定の上位であり、「噴火／噴水」「鉄船／巨大な壁」はわかりやすさの判定の上位でもある表現であり、比喩性とわかりやすさと具体化には関連が存在することが予測される。

表 21 に具体化の評定が下位の事例を示す。「口にする」が 4 例存在する。「X にする」形式が換喩的であり、積極的に表現対象を他の事物・事柄に見立てるものではないために具体化ではないと判断されたと推測される。「思い出したように」はある人物の話し方に対する表現主体の

表 19 擬人化の評定 (下位事例)

サンプル ID	比喩事例	擬人化
PB35.00013	「ここまでが標準仕様で、あとはオプションとなります」という言葉をよく耳にしませんか？	0.44
PM26.00004	最終的には円盤状病巣といわれる厚い線維性瘢痕組織となり網膜機能は失われる。	0.44
OW6X.00073	タイ及びその周辺国への薬物分析技術の移転等を柱とする「薬物対策地域協力プロジェクト」に、	0.48
PB1n.00024	午前十時四十分、議員バッジを 1 年半ぶりに胸にする加藤氏は言った。	0.50
PN3a.00017	火口だった場所は、すり鉢のように深くえぐれている。	0.50

表 20 具体化の評定 (上位事例)

サンプル ID	比喩事例	具体化
PM11.00260	化粧品が天才的にヘタでピカソみたいなメークになってる女や、	4.45
PB1n.00024	溶岩は、粘りが少ないため、噴火すると噴水のようにふきあがる。	4.35
PB39.00023	泳ぎつかれたら一本の棒みたいに海水に浮かび、	4.35
PB39.00009	大きな鉄船は巨大な壁のように目の前をふさいでいる。	4.23
OC09.03043	内出血の場合だとペンキをペタッと塗ったように真っ赤になります。	4.19

表 21 具体化の評定 (下位事例)

サンプル ID	比喩事例	具体化
PB39.00010	漸く一言口にして、彼は絶句した。	1.09
PM12.00014	ダグは具体的なことを口にしなかった。	1.20
PM21.00296	杉田がたまに思い出したようにそんな話をしても、(中略) 実際にはまともに取り合わなくなっていた。	1.25
PM21.00296	でも今は「やらせて下さい」と、無理にでも口になっている。	1.26
OC09.04669	しかも許すって言ったら 2 度と口にしないのが鉄則です。	1.34

印象を示すものであり、具体化しようとする例とは判断されない。

#### 4.2.2 クラウドソーシングによる評定値の統計分析

以下では得られたクラウドソーシングデータに関して、例文のバイアスおよび実験協力者のバイアスを考慮した統計分析を行う。rstan (Stan Development Team 2016) により、レジスタによる効果を分析するモデル (1) とレジスタ・種別による効果を分析するモデル (2) の線形混合モデルを構築し、ベイズ推定する。具体的には、比喩性・新奇性・わかりやすさ・擬人化・具体化の各パラメータを、提示文と実験協力者をランダム傾き (変量効果) としたうえで、レジスタと種別を傾き (固定効果) として定義した次式によりモデル化する。

$$y \sim \text{Normal}(\mu, \sigma),$$

$$(1) \quad \mu \leftarrow \alpha + \beta_{\text{レジスタ}}^{OC} * \chi_{OC}(x) + \beta_{\text{レジスタ}}^{OW} * \chi_{OW}(x) + \beta_{\text{レジスタ}}^{OY} * \chi_{OY}(x) + \beta_{\text{レジスタ}}^{PB} * \chi_{PB}(x) + \beta_{\text{レジスタ}}^{PM} * \chi_{PM}(x) + \beta_{\text{レジスタ}}^{PN} * \chi_{PN}(x) + \gamma_{\text{例文}}^* + \gamma_{\text{実験協力者}}^*$$

$$(2) \quad \mu \leftarrow \alpha + \beta_{\text{レジスタ}}^{OC} * \chi_{OC}(x) + \beta_{\text{レジスタ}}^{OW} * \chi_{OW}(x) + \beta_{\text{レジスタ}}^{OY} * \chi_{OY}(x) + \beta_{\text{レジスタ}}^{PB} * \chi_{PB}(x) + \beta_{\text{レジスタ}}^{PM} * \chi_{PM}(x) + \beta_{\text{レジスタ}}^{PN} * \chi_{PN}(x) + \beta_{\text{種別}}^{\text{擬人}} * x_{\text{種別}}^{\text{擬人}} + \beta_{\text{種別}}^{\text{擬生}} * x_{\text{種別}}^{\text{擬生}} + \beta_{\text{種別}}^{\text{擬物}} * x_{\text{種別}}^{\text{擬物}} + \beta_{\text{種別}}^{\text{具象}} * x_{\text{種別}}^{\text{具象}} + \beta_{\text{種別}}^{\text{抽象}} * x_{\text{種別}}^{\text{抽象}} + \beta_{\text{種別}}^{\text{転換}} * x_{\text{種別}}^{\text{転換}} + \beta_{\text{種別}}^{\text{換喩}} * x_{\text{種別}}^{\text{換喩}} + \beta_{\text{種別}}^{\text{提喩}} * x_{\text{種別}}^{\text{提喩}} + \beta_{\text{種別}}^{\text{文脈}} * x_{\text{種別}}^{\text{文脈}} + \beta_{\text{種別}}^{\text{慣用}} * x_{\text{種別}}^{\text{慣用}} + \beta_{\text{種別}}^{\text{その他}} * x_{\text{種別}}^{\text{その他}} + \gamma_{\text{例文}}^* + \gamma_{\text{実験協力者}}^*$$

ここで、 $y$  は比喩性・新奇性・わかりやすさ・擬人化・具体化の各パラメータとする (評定値ごとに推定を行う)。Normal は平均  $\mu$  標準偏差  $\sigma$  の正規分布とし、切片  $\alpha$  とレジスタ・種別に対応する傾き ( $\beta_{\text{レジスタ}}^*$ ,  $\beta_{\text{種別}}^*$ ) と例文・実験協力者のバイアス ( $\gamma_{\text{例文}}^*$ ,  $\gamma_{\text{実験協力者}}^*$ ) の線形式<sup>26</sup>である。各変数の事前分布は標準正規分布とし、ハイパーパラメータ  $\sigma$  も含めて、ベイズ推定 (3 chains, warmup 100, 1,000 iteration) を行う。

有効データポイント数は 27,072、提示文数 820、実験協力者数 1,164 であった。Gelman-Rubin 統計量 (Rhat) は 1.1 未満となっていないときに、連鎖が定常状態に収束していると判断した。本分析では、構築したいずれのモデルも収束した。なお、同等のモデルを頻度主義的な線形混合モデル (lmer) により推定したところ収束しなかった。統計的な検討は事後平均 (mean) が事後標準偏差 (sd) の 2 倍以上の差がある場合に、強い証拠がある (以下の表中太字で示す) と判断し、事後偏差の差がある場合に弱い証拠があると判断する<sup>27</sup>。表 22, 23 にレジスタによる効果を分析するモデル (1) とレジスタ・種別による効果を分析するモデル (2) の結果を示す。

<sup>26</sup> ここで  $\chi_A(x) := \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$  (そのレジスタか否かの指示関数)、 $x_{\text{種別}}^A := x$  中の種別 A の結合の数とする。

<sup>27</sup> 本分析は帰無仮説に基づかないために、帰無仮説を前提とする有意差の議論は行わない。ベイズ推定に基づく同様の用語として「強い証拠 (strong evidence)」「弱い証拠 (weak evidence)」を用いる。

表 22 統計分析結果：レジスタによる効果を分析するモデル (1)

Parameter	比喩性		新奇性		わかりやすさ		擬人化		具体化	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
$\alpha$	2.45	0.39	1.88	0.40	2.97	0.40	1.40	0.39	2.76	0.44
$\beta_{レジスタ}^{OC}$	-0.13	0.44	-0.24	0.43	0.21	0.42	-0.02	0.44	-0.12	0.47
$\beta_{レジスタ}^{OW}$	-0.75	0.39	-0.49	0.42	-0.11	0.42	-0.18	0.39	-0.33	0.46
$\beta_{レジスタ}^{OY}$	0.32	0.41	0.12	0.42	-0.03	0.40	0.09	0.41	0.00	0.45
$\beta_{レジスタ}^{PB}$	0.47	0.40	0.10	0.41	0.08	0.40	0.12	0.38	0.17	0.45
$\beta_{レジスタ}^{PM}$	0.32	0.39	0.07	0.41	0.05	0.40	-0.00	0.40	0.11	0.45
$\beta_{レジスタ}^{PN}$	0.02	0.39	-0.11	0.41	0.13	0.41	-0.01	0.40	0.04	0.45

表 23 統計分析結果：レジスタ・種別による効果を分析するモデル (2)

Parameter	比喩性		新奇性		わかりやすさ		擬人化		具体化	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
$\alpha$	2.38	0.42	1.81	0.44	2.90	0.42	1.27	0.42	2.76	0.44
$\beta_{レジスタ}^{OC}$	0.04	0.46	-0.04	0.46	0.28	0.44	0.11	0.41	-0.02	0.47
$\beta_{レジスタ}^{OW}$	-0.45	0.45	-0.27	0.45	-0.05	0.43	0.02	0.43	-0.19	0.42
$\beta_{レジスタ}^{OY}$	0.23	0.43	0.17	0.43	-0.04	0.42	0.18	0.41	-0.08	0.45
$\beta_{レジスタ}^{PB}$	0.22	0.42	0.10	0.43	0.05	0.42	0.12	0.41	-0.01	0.43
$\beta_{レジスタ}^{PM}$	0.15	0.42	0.12	0.43	0.03	0.42	0.08	0.41	-0.03	0.44
$\beta_{レジスタ}^{PN}$	-0.02	0.42	-0.04	0.44	0.14	0.42	0.08	0.41	-0.04	0.43
$\beta_{種別}^{擬人}$	<b>0.37</b>	0.15	<b>0.30</b>	0.15	0.02	0.15	<b>0.75</b>	0.14	0.11	0.15
$\beta_{種別}^{擬生}$	<b>0.67</b>	0.14	0.23	0.15	0.03	0.15	<b>0.59</b>	0.16	0.29	0.16
$\beta_{種別}^{擬物}$	<b>0.46</b>	0.15	0.25	0.14	0.10	0.15	<b>0.41</b>	0.14	0.19	0.16
$\beta_{種別}^{具象}$	0.12	1.06	0.04	1.04	-0.15	0.93	0.05	0.95	-0.05	1.03
$\beta_{種別}^{抽象}$	0.46	0.51	0.36	0.52	-0.17	0.49	0.13	0.49	0.34	0.51
$\beta_{種別}^{転換}$	<b>0.48</b>	0.08	0.08	0.08	0.15	0.08	0.03	0.08	<b>0.32</b>	0.08
$\beta_{種別}^{換喩}$	-0.32	0.16	0.01	0.15	-0.07	0.15	-0.00	0.15	-0.24	0.15
$\beta_{種別}^{提喩}$	<b>0.32</b>	0.10	0.08	0.10	0.08	0.11	-0.02	0.10	0.29	0.10
$\beta_{種別}^{文脈}$	-0.17	0.33	0.10	0.31	-0.15	0.31	0.15	0.32	-0.23	-0.31
$\beta_{種別}^{慣用}$	-0.72	0.10	-0.36	0.10	0.05	0.10	-0.20	0.11	-0.45	0.10
$\beta_{種別}^{その他}$	0.21	0.14	-0.03	0.15	0.06	0.13	0.01	0.14	0.22	0.14

まず、比喩性の結果について確認する。レジスタにおいては、レジスタの標準偏差が大きい  
ため、白書 (OW) の傾きが小さく、書籍 (PB) の傾きが大きいという弱い証拠のみが確認され  
た。種別においては、擬人・擬生・擬物・転換・提喩の傾きが大きく、換喩・慣用の傾きが小  
さいという強い証拠が確認された。また、具象・抽象は標準偏差が大きい傾向が見られた。

活喩（擬人化・擬生化）の場合、クラウドソーシングによる擬人化の評定とも関係するが、比喩性が高いと認識される傾向にあった。擬人化と対照的な擬物化も高く評定されており、人に関わる比喩表現は、一般に比喩性が読み取られやすいといえよう。しかし、具象化の比喩性は様々に判定されている。例を見ると、「研究の基盤」「時の流れ」「(以下の点を)柱とする・軸とする」のように、慣用的に用いられている表現が多く、これらは新奇性の評定とともに比喩性の判定が低いためである。具象化は多くの結合類型例と頻度を有する分類であるからこそ、一般的な表現であって比喩性を感じにくいともいえる。また、小規模な転換（その他の転換）は、比喩性の高い傾向にあるが、「クレーターのようなくぼみ」のような例示に近い例や、「自転車がバイクに似ている」のように喩辞と被喩辞の類似性が高いと考えられる結合例も含み、例によっては比喩性の判定に差が生じたようである。換喩は、「耳にする」「口にする」などで比喩性が低く判定される傾向があり、比喩表現と認識されにくい例が多いようである。文脈比喩については、解釈に長い文脈を要する例もあり、実験で表示された該当部分では比喩性が読み取りにくかった影響も考えられる。

次に、新奇性の結果について確認する。レジスタにおいては、白書(OW)の傾きが小さいという弱い証拠のみが確認された。種別においては擬人の傾きが大きく、慣用の傾きが小さいという強い証拠が確認された。また擬生・擬物・転換の傾きが大きいという弱い証拠が確認された。

比喩性が認識されやすい人を用いた表現（擬人・擬物）で新奇性が高く評定される傾向にある。慣用はその定義から新奇性が低いのは当然の結果であると言える。また、換喩は相対的に傾きが小さい（事後平均 0.01）傾向にある。換喩は一般に用いられている慣用的な表現が多い（「慣用」と重なりやすい）ため、新奇な印象は得にくいものと考えられる。小規模な転換については、指標を必要として比喩性の把握を促すような表現では高く判定されるが、比喩性が低いと評定された喩辞と被喩辞の類似性が低い表現ではみとめられない傾向にあった。

わかりやすさにおいては、レジスタ間に特段の傾向は見られなかった。種別間は、転換の傾きが大きいという弱い証拠が見られた。

指標を含む比喩用例は、概ねいくぶんわかりやすいという印象で読まれているようである。分類中、慣用表現の評定が他分類よりも僅かに高く、提喩や活喩など、一般に目にしやすいと考えられる表現ほど、わかりやすいと判断される傾向がみられる。比喩表現がわかりやすさに直結するというよりも、新奇でない目にしやすい表現がわかりやすいという判断につながっている可能性もある。

擬人化の結果について検討する。まずレジスタにおいては特段の傾向は見られなかった。種別においては、擬人・擬生・擬物の傾きが大きいという強い証拠が見られた。また、慣用の傾きが小さいという弱い証拠が見られた。

擬人化の用例で高い傾向がみられることは、専門家の判断と一般の判断にそこまで揺れがないということでもある。しかし、擬生化に分類された用例でも高く評定されている他、擬物化

でも高めの評定となっている。喩辞が人に限らないとしても概ね生命に関する表現が含まれ、結合としてわかりやすいためであろう（わかりやすさの結果を参照）。また、活喩については、一般には、人と生命体（人には限定できない）を厳密に分類せず、生命活動であるという点において読み手自らの身体性に鑑みるといえる。擬物化は、擬人化とは反対に人が被喩辞として物に喩えられる例であるが、擬物化も人を用いた比喩表現であるとの判断で、人に喩えた例と判断されたようである。なお、擬人化・擬生化・擬物化については、「子どもたちは魚のように水と友だちだった」「次郎は舟と一体になる」のような重複ラベル用例も含んでいる。

具体化においては、レジスタ間の差異に特段の傾向は見られなかった。種別間では、転換の傾きが大きく、慣用の傾きが小さいという強い証拠が見られた。また擬生・擬物・提喩・その他の傾きが大きく、換喩の傾きが小さいという弱い証拠が見られた。

人に関わる用例において高く評定される傾向は、生命活動に関わる比喩表現が身体性としてわかりやすく、具体的であると認識されるためであろう。また、標準偏差が大きいものの具象化の用例は、具象化と判断されにくい傾向を示す結果となった。具象化の用例は、比喩性においても低い評定傾向であったことから、概念メタファー (Lakoff and Johnson 1980) として一般的、かつ慣用的な用例が多いため、言語形式が比喩表現の把握を促していても、指標まで含めたイディオム的な認識となり、具象化しているとあらためて認識されない可能性が高い。

## 5 おわりに

本研究では、比喩の指標の可能性のある語句を手掛かりとし、BCCWJから比喩用例を収集した。手掛かりには359種類の指標とそれらの語義タグも使い、広く指標となり得る語句と語義を含む網羅的な指標比喩の可能性を有する用例を収集している。指標比喩の可能性のある97,118用例から人手で822件を抽出し、比喩情報のアノテーションを行った。

収集した用例と付与した情報により、指標比喩の特徴を確認した。種々の指標比喩を収集する試みにより、たとえばいわゆる直喩において一般的な指標と考えられる「よう」が本データでは43%にとどまるなど、指標要素が多様であることをはじめ、指標比喩タイプの分布傾向を見ることができた。さらに、比喩種別を分類すると、指標比喩においては、結合比喩に少ない「その他の転換」とまとめた小規模な転換の頻度が明らかに高いとわかった。喩辞と被喩辞がかけ離れている大きなカテゴリーの転換（擬人化や具象化など）がある場合には、比喩性があきらかで指標が不要な場合が多く、小規模な転換であるゆえに言語形式で比喩性を示す傾向にあるといえる。指標の有無は、比喩表現を構成する要素の概念マッピングの幅に関わっている可能性が考えられる。今後、比喩表現の指標の有無による差異についても調査を進めたい。

また、クラウドソーシングによる比喩種別の分類から、観点の関係性が読み取れた。擬人化や擬生化、擬物化の評定結果など、専門知識を有する作業者と一般的な読み手間における認識差

の影響が現れた観点もあるが、本調査で一般的な比喩認識が用例に付与されたことにより、明らかになった差であるともいえる。但し、複数の表現（比喩的な要素の結合）が互いに影響しあっている可能性も考えられるため、今後は表現別に観定の評定を行う他、影響関係についても評定実験を考えている。

伊藤 (2014) は、比喩表現コーパスに望むこととして以下をあげている：

- 一つのテキストを通して比喩のタグ付けをすること
- データとした談話やテキストは、全文を閲覧可能であること
- データのメタ情報にアクセス可能なこと
- 判断の揺れるものについては、できるだけ比喩であると認めること
- オープンであること
- 比喩の判断基準を明示すること

我々のデータは以上のいずれも満たす。さらに、中村 (1977) の文学作品中心の比喩用例データベースと異なり、Yahoo!知恵袋・白書・Yahoo!ブログ・書籍・雑誌・新聞の6レジスタを対象として収集を行った。本データにより、文学作品以外にも比喩が多用されていることが確認できた。

現在、分類語彙表番号を意味情報として用いた、結合比喩（中村の第2類）パターンに基づく、比喩表現用例の収集も進めている (Kato and Asahara 2019)。また、語義や結合の認識の効率化のために、文脈化単語埋め込み (Peters, Neumann, Iyyer, Gardner, Clark, Lee, and Zettlemoyer 2018) の利用を検討している。『国語研日本語ウェブコーパス』 (Asahara, Maekawa, Imada, Kato, and Konishi 2014) から学習した自然言語処理の BERT (Devlin, Chang, Lee, and Toutanova 2019) モデルである NWJC-BERT (浅原, 西内, 加藤 2020) を構築し、BCCWJ に文脈化単語埋め込み情報を付与した (浅原, 加藤 2020)。本研究で得られたデータと重ね合わせることで比喩情報付き均衡コーパスのさらなる整備を進めたい。

## 謝 辞

本研究は、The International Cognitive Linguistic Conference (ICLC15) の発表 “Collecting Figurative Expressions Using Indicators and a Semantic Tagged Japanese Corpus” に基づくもの、追加実験を行い、加筆修正をしたものです。本研究は、国立国語研究所コーパス開発センター共同研究プロジェクト「コーパスアノテーションの拡張・統合・自動化に関する基礎研究」によるものです。本研究の一部は JSPS 科研費 挑戦的研究 (萌芽) 18K18519, 基盤研究 (A) 17H00917, 新学術領域研究 18H05521, 基盤研究 (C) 18K00634, 19K00591 の助成を受けたものです。

## 参考文献

- 浅原正幸, 加藤祥 (2020). BERTed-BCCWJ: 多層文脈化単語埋め込み情報を付与した『現代日本語書き言葉均衡コーパス』データ. 言語処理学会第 26 回年次大会発表論文集, pp. 161–164.
- Asahara, M., Maekawa, K., Imada, M., Kato, S., and Konishi, H. (2014). “Archiving and Analysing Techniques of the Ultra-large-scale Web-based Corpus Project of NINJAL, Japan.” *Alexandria: The Journal of National and International Library and Information Issues*, **25** (1–2), pp. 129–148.
- 浅原正幸, 西内沙恵, 加藤祥 (2020). NWJC-BERT: 多義語に対するヒトと文脈化単語埋め込みの類似性判断の対照分析. 言語処理学会第 26 回年次大会発表論文集, pp. 961–964.
- Bowdle, B. F. and Gentner, D. (2005). “The Career of Metaphor.” *Psychological Review*, **112**, pp. 193–216.
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.” In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pp. 4171–4186.
- 古谷望, 寺井あすか (2018). 参照表現利用に対する比喩の効果の検討. 言語処理学会第 24 回年次大会発表論文集, pp. 200–203.
- 榛谷泰明 (1988). レトリカー比喩表現辞典. 白水社.
- 伊藤薫 (2014). 比喩表現コーパスの構築と問題点—言語学の立場から—. 言語処理学会第 20 回年次大会発表論文集, pp. 149–152.
- 神崎享子 (2013). 『複合動詞レキシコン』 ver.1—形態的・統語的・意味的情報付与—. 言語処理学会第 19 回年次大会発表論文集, pp. 761–764.
- Kato, S. and Asahara, M. (2019). “Exploring Metaphorical Expressions in Japanese Newspaper-Article Corpora.” In *International Cognitive Linguistics Conference 15*.
- 加藤祥, 浅原正幸, 山崎誠 (2019a). 『現代日本語書き言葉均衡コーパス』新聞・書籍・雑誌データの助動詞に対する用法情報付与. 日本語学会 2019 年度春季大会予稿集, pp. 169–174.
- 加藤祥, 浅原正幸, 山崎誠 (2019b). 分類語彙表番号を付与した『現代日本語書き言葉均衡コーパス』の書籍・新聞・雑誌データ. 日本語の研究, **15** (2), pp. 134–141.
- 菊地礼, 加藤祥, 浅原正幸 (2018). 「感じる」を指標とするメタファー用例の収集とその分析. 日本語用論学会メタファー研究会 2-Day シンポジウム「身体性」.
- 国立国語研究所 (編) (2004). 国立国語研究所資料集 14 『分類語彙表—増補改訂版—』. 大日本図書.
- 近藤明日子, 田中牧郎 (2020). 「分類語彙表番号—UniDic 語彙素番号対応表」の構築. 国立国

語研究所論集, pp. 77–91.

楠見孝 (1995). 比喩の処理過程と意味構造. 風間書房.

Lakoff, G. and Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press.

レイコフ・ジョージ, ジョンソン・マーク (1986). レトリックと人生. 渡部昇一, 楠瀬淳三, 下谷和幸 (訳). 大修館書店.

Maekawa, K., Yamazaki, M., Ogiso, T., Maruyama, T., Ogura, H., Kashino, W., Koiso, H., Yamaguchi, M., Tanaka, M., and Den, Y. (2014). “Balanced Corpus of Contemporary Written Japanese.” *Language Resources and Evaluation*, **48**, pp. 345–371.

宮島達夫 (1972). 動詞の意味・用法の記述的研究. 秀英出版.

宮澤彬, 吉田奈央, 宮尾祐介 (2016). 日本語メタファーコーパス作成のためのガイドライン. 言語処理学会第22回年次大会発表論文集, pp. 150–153.

Nakamoto, K. (2003). “Semantic Priming Effect of Metaphor Constituent Terms.” *Perceptual and Motor Skills*, **96**, pp. 33–42.

中本敬子, 楠見孝 (2004). 比喩材料文の心理的特性と分類—基準表作成の試み—. *読書科学*, **48** (1), pp. 1–10.

中村明 (1977). 比喩表現の理論と分類. 国立国語研究所報告 57. 秀英出版.

中村明 (1995). 比喩表現辞典. 角川書店.

中村明 (2007). 日本語の文体・レトリック辞典. 東京堂出版.

野内良三 (1998). 日本語の文体・レトリック辞典. 国書刊行会.

野内良三 (2005). 日本語修辭辞典. 国書刊行会.

Ortony, A., Vondruska, R. J., Foss, M. A., and Jones, L. E. (1985). “Salience, Similes, and the Asymmetry of Similarity.” *Journal of Memory and Language*, **24**, pp. 569–594.

Peters, M., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., and Zettlemoyer, L. (2018). “Deep Contextualized Word Representations.” In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pp. 2227–2237.

Pragglejaz Group (2007). “MIP: A Method for Identifying Metaphorically Used Words in Discourse.” *Metaphor and Symbol*, **22** (1), pp. 1–39.

岡隆之介, 大島裕明, 楠見孝 (2019). 比喩研究のための直喩刺激—解釈セット作成および妥当性の検討. *心理学研究*, **90** (1), pp. 53–62.

佐藤信夫, 佐々木健一, 松尾大 (2006). レトリック事典. 大修館書店.

Spanger, P. M., Yasuhara, M., Iida, R., and Tokunaga, T. (2012). “REX-J: Japanese Referring Expression Corpus of Situated Dialogs.” *Language Resources Evaluation*, **46** (3), pp. 461–491.

- Stan Development Team (2016). “RStan: the R interface to Stan.” R package version 2.14.1.
- Steen, G. J., Dorst, A. G., Herrmann, J. B., Kaal, A., Krennmayr, T., and Pasma, T. (2010). *A Method for Linguistic Metaphor Identification*. John Benjamins Publishing.
- 須堯大喜, 寺井あすか (2019). 係り受け解析を用いた比喩生成支援システム. 情報処理学会研究報告 CH-119 No.1, pp. 1-6.
- 平知宏, 中本敬子, 楠見孝 (2007). 比喩理解における親しみやすさと解釈の多様性. 認知科学, **14** (3), pp. 323-338.
- Utsumi, A. (2007). “Interpretive Diversity Explains Metaphor-simile Distinction.” *Metaphor and Symbol*, **22**, pp. 291-312.
- 山梨正明 (1988). 比喩と理解. 東京大学出版会.
- 山崎誠, 柏野和佳子 (2017). 『分類語彙表』の多義語に対する代表義情報アノテーション. 言語処理学会 第 23 回年次大会 発表論文集, pp. 302-305.

## 略歴

**加藤 祥**：2011 年神戸大学人文学研究科博士後期課程修了。2012 年より国立国語研究所コーパス開発センタープロジェクト PD フェロー。同プロジェクト非常勤研究員。2020 年より目白大学外国語学部専任講師。博士（文学）。日本語学会，日本認知言語学会，日本認知科学会各会員。

**菊地 礼**：2018 年中央大学文学研究科博士課程前期修了。同年より同大学博士課程後期に在学。

**浅原 正幸**：2003 年奈良先端科学技術大学院大学情報科学研究博士後期課程修了。2004 年より同大学助教。2012 年より人間文化研究機構国立国語研究所コーパス開発センター特任准教授。2019 年より同教授。博士（工学）。言語処理学会，日本言語学会，日本語学会各会員。

(2020 年 4 月 30 日 受付)

(2020 年 7 月 31 日 再受付)

(2020 年 9 月 10 日 採録)