

# 絵本の言葉についての一考察 ——「ぐりとぐら」シリーズを資料として——

## A Study on the Vocabulary of Picture Books: “GURI and GURA” Series as Text

平林 一利

Kazutoshi Hirabayashi

### はじめに

子どもの言葉の発達において、児童文化財である絵本との関わりは重要であり、必要とされる<sup>注1)</sup>。絵本が子どもの言葉の発達にどのように関わり、子どもの言葉の育みにどのように役立つのかを考察するためには、幼児絵本にどのような言葉が用いられているのかを調査分析することが必要である。

本稿では、その取り組みの一環として幼児絵本に見られる語彙の数量的・意味的な調査を試みる。なお、このような絵本を対象とした、中曾根・川又（1994）による大規模な語彙調査や、平・藤田・小林（2012）、藤田・小林・奥村（2017）による電子計算機を用いた大規模データの形態素解析調査がおこなわれコーパスも作成され評価されている。

### 調査資料

幼児用の絵本はブックスタート<sup>注2)</sup>以後注目され、近年では多くの絵本が出版されている。そのため調査対象とすべき絵本は多く出版されている。本稿では前稿<sup>注3)</sup>を踏まえ、長期にわたり子どもたちに読み継がれている中川李枝子（作）・大村百合子（絵）「ぐりとぐら」シリーズの7冊を調査対象とする<sup>注4)</sup>。書名と初版年は以下のとおりである。今回使用したテキストの出版年と刷を（ ）内に示す。

『ぐりとぐら』（『母の友』（1963）所収）	1963年初版	（2021年	236刷）
『ぐりとぐらのおきゃくさま』	1966年初版	（2008年	128刷）
『ぐりとぐらのかいすいよく』	1976年初版	（1997年	60刷）
『ぐりとぐらのえんそく』	1979年初版	（2007年	93刷）
『ぐりとぐらとくるりくら』	1987年初版	（2003年	45刷）

- 『ぐりとぐらとすみれちゃん』 2000年初版 (2019年 44刷)  
『ぐりとぐらのおおそうじ』 2002年初版 (2008年 25刷)

これらを資料とするのは、1963年の出版以来継続的に読み継がれシリーズ化され、いわゆる名作とされる作品であり、保育者も含め、多くの子どもたちに読み語られるものであるため、絵本の語彙研究の基礎的な資料として適切であると考ええる。

### 調査方法

「ぐりとぐら」シリーズの7冊を電子データ化し国立国語研究所が公開している、形態素解析アプリケーション「Web茶まめ」<sup>注5)</sup>を用いて形態素解析をおこなった<sup>注6)</sup>。

調査対象とした資料の本文は平仮名と片仮名のみであるので全文を漢字仮名交じり文に変換し、「Web茶まめ」の解析項目を「現代語辞書」「語彙素」「語彙素読み」「品器大分類」「語種」と指定し出力した。

「Web茶まめ」から出力された解析データから、単語（形態素）の切れ目、「語彙」「語彙素読み」「品詞」「語種」の情報を確認し、形態素より長い単位として扱うべきと考えられる例や、誤解析されている箇所について手作業で修正を施した。なお、句読点、くくり記号などの記号とされるものは、言葉ではないとし、今回の調査対象から除外した。

#### 解析について

形態素よりも長い単位での認定の例と品詞の修正の例を一部示す。

#### 形態素より長い単位での認定

「Web茶まめ」の解析はかなり細かく、接頭辞・接尾辞（接辞的）についても解析されるため、以下のように、これらを含む語を一つの言葉として認定した。また、使用実態にあった単位に修正し、品詞についても修正をおこなった。

- ・あわだて き→ 泡立て器（あわだてき） ・一 度→ 一度（いちど）
- ・御 砂糖→ お砂糖（おさとう） ・お じい（爺）さん→ お爺さん（おじいさん）
- ・波打ち 際→ 波打ち際（なみうちぎわ） ・二 匹→ 二匹（にひき）
- ・目覚まし 時計→ 目覚まし時計（めざましどけい） ・雪 達磨→ 雪達磨（ゆきだるま）
- ・雪 合戦→ 雪合戦（雪合戦） ・お 土産→ お土産

#### 品詞の修正について

- ・おまけ に（名詞・助詞）→ 接続詞 ・くるり と（名詞・助詞）→ 副詞
- ・ところ が（名詞・助詞）→ 接続詞 ・ところ で（名詞・助詞）→ 接続詞
- ・そう と も（副詞・助詞・助詞）→ 感動詞

などである<sup>注7)</sup>。

## 「ぐりとぐら」シリーズに見られる語の量的な性質

修正を施したデータをもとに、調査資料から得た語彙を集計し述べ語数4735語、異なり語数928例を得た。

## 語種の構成比率

「ぐりとぐら」シリーズに見られる用例の総数を語種別に表1として示す。なお、頻出する「ぐり」「ぐら」は和語としている。

表1 語種の異なり語数と述べ語数

	和語	漢語	外来語	混種語	計
異なり語数	715	105	55	53	928
述べ語数	4300	207	122	106	4735

語種別にみると、異なり語数・述べ語数ともに和語が多数を占めていることがわかる。日本語を資料とした全数調査のばあい、その特徴として助詞・助動詞が多数を占める。そのため解析結果に偏りがおこることが予想されるので、助詞・助動詞を除き自立語のみで集計したものが表2である。

表2 自立語の語種の異なり語数と述べ語数

	和語	漢語	外来語	混種語	計
異なり語数	669	105	55	53	882
述べ語数	2065	207	122	106	2498

助詞・助動詞は基本的に和語のため、表2では和語の語数が減少するが、和語の数値が高いことに変化はない。日本語の大規模データを用いた語種の割合については『現代雑誌90種の用語用字第三分冊』<sup>注8)</sup>のものが参考にされる。表3として下記に示す。

表3 現代雑誌90種の用語用字における語種分布（% 四捨五入）

	和語	漢語	外来語	混種語	計
異なり語数	36.7%	47.5%	9.8%	6.0%	100.0%
述べ語数	53.9%	41.3%	2.9%	1.9%	100.0%

次に、表2に示した自立語の語種別の割合を表4として示す。

表4 語種分布 (% 四捨五入)

	和語	漢語	外来語	混種語	計
異なり語数	75.9%	11.9%	6.2%	6.0%	100.0%
述べ語数	82.6%	8.3%	4.9%	4.2%	100.0%

一般的な文章に見られる語種の割合は、述べ語数では和語の数値が高く、異なり語数では漢語の数値が高くなるが、表4では、漢語の異なり語数が若干増加するが、語種の割合に極端な変化はみられない。

また『計量国語学事典 新装版』には「児童読み物の調査 (異なり語：3767語)」<sup>注9)</sup>の調査結果が示されているが、

和語64.0% 漢語29.7% 外来語4.0% 混種語2.3%

となっている。今回調査した幼児絵本である「ぐりとぐら」シリーズは、児童の読み物と比べても和語の割合が高いことが見てとれ、幼児絵本の特徴と考えられる。

#### 「ぐりとぐら」シリーズに見られる語彙の量的構造

「ぐりとぐらシリーズ」に見られる言葉を数量的に考察する。先に示した表2の異なり語数、述べ語数を元に、特に頻出する「ぐり」「ぐら」の2語259例を除き、図1と図2を作成した。

図1は、異なり語数と述べ語数をもとに度数分布をグラフにまとめたものであり、図2は、使用頻度の高いものを横軸に示し、使用頻度の高いものから順位をとり、縦軸に累積使用率を示したものである。

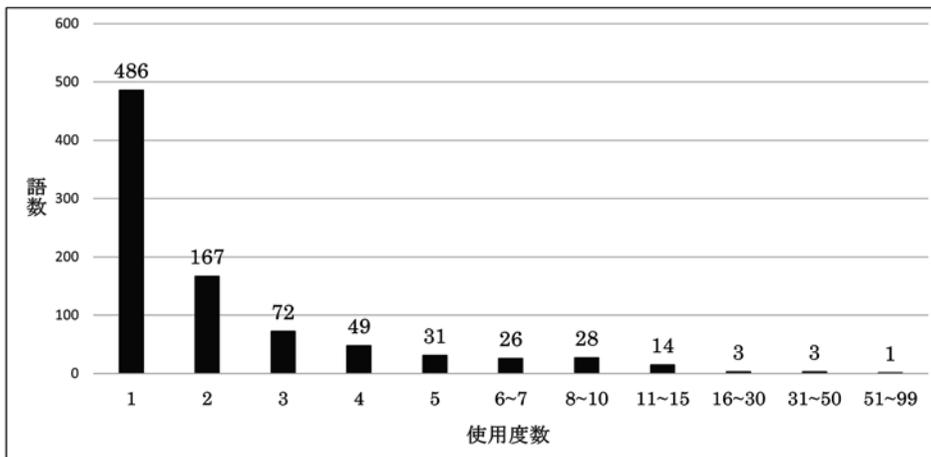


図1 語彙の度数分布

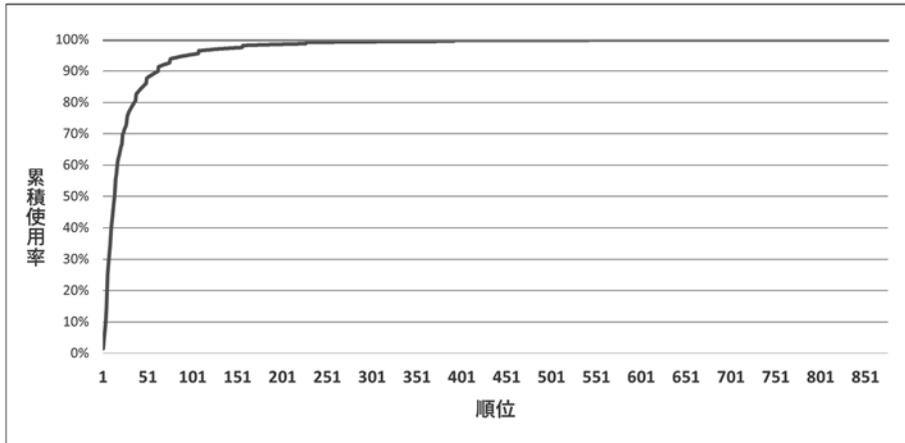


図2 語の累積比率

図1は、度数1の語が多く使われていることを示し、度数が上がるにつれ語数が減少していることが見てとれる。図2は、上位50語（5.6%）で資料全体の約80%を占め、上位100語（11.3%）で95%程度を占めていることを示す。図1、図2の結果は、自立語では、使用頻度の高い一部の語だけで語彙全体のかなりの部分を占めていることがわかる。なお、このような語彙の量的な構造は、L字曲線として示され、一般的な日本語の文章において共通するものである<sup>注10)</sup>。

### 語彙のランク

高頻度の語彙と低頻度の語彙にはどのような語が所属しているのかを確認する。このことを考えるために、図1のもととなった語別頻度表から語彙使用ランクを田中（2016）<sup>注11)</sup>を参考にランク別に5段階に分けた<sup>注12)</sup>。日本語を対象としたこのような調査の場合、通常ランクⅠは、助詞・助動詞がほぼ100%を占めることになるので調査対象から除き、「ぐり」と「ぐら」を含めた自立語（自立語相当語含む）の使用ランクから考察する。ランク分けは表5のとおりである。

表5 度数による語彙のランク

	I	II	III	IV	V
度数（回）	29～	11～18	4～9	2～3	1
語数（語）	5	17	132	239	486

「Web茶まめ」で形態素解析した品詞分類のデータをもとに、品詞としての類似性の高いものを統合し、a 体言（名詞・代名詞）、b 用言（動詞・形容詞・形状詞）、c 副詞・連体詞、d 接続詞・感

動詞の4種にまとめ、それぞれの割合を表6として示す。

語彙ランクと品詞

表6 語彙ランクと品詞 (% 四捨五入)

ランク	a	b	c	d
I	54.3%	45.7%	0%	0%
II	51.0%	46.5%	3.5%	0%
III	56.0%	29.0%	9.7%	5.3%
IV	53.0%	29.7%	9.4%	7.9%
V	50.8%	35.1%	9.7%	4.4%

a 体言 (名詞・代名詞)    b 用言 (動詞・形容詞・形状詞)    c 副詞・連体詞    d 接続詞・感動詞

高頻度のランク I に含まれるのは a と b である。次に高頻度のランク II の語彙は a、b に加え c 副詞・連体詞が見られるが僅かであり、a と b が高い比率を占めていることにはかわりはない。ランク III～V では a～d すべてそろろうが、a の比率が高いまま、b が減少しているのに対して c と d が増加するが、d はランク V で減少する。次に実際にどのような言葉が高頻度語として表れるのか確認するためにランク I～III の語を品詞別に表7に示す<sup>注13)</sup>。

表7 高頻度語の語彙

	品詞	ランク I	ランク II	ランク III
a	名詞 代名詞	ぐり、ぐら、	家 (いえ)、海坊主、カボチャ、くるりくら、こと、これ、すみれちゃん、手、のねずみ、僕たち、帽子	朝ご飯、丘、上、お日様、籠、カステラ、糸糸、熊、ここ、そこ、それ、とき、どこ、二匹、野原、僕、僕ら
	動詞	言う、行く、居る、来る、する	ある、作る、叩く、なる、入る、持つ、見る	開ける、急ぐ、歌う、踊る、思う、食べる、出来る、出す、飛ぶ、やる
	形容詞 形状詞		大きい、無い、良い	赤い、青い、美味しい、丸い、 こんなに、好き
c	副詞			いっぱい、ずっと、そよそよ、どう、とても、ふわふわ、ほんと、もう
	連体詞		この	大きな、その
d	接続詞			
	感動詞			ありがとう、うん、ほら、やあ

ランクⅠ、Ⅱに挙がる名詞は、それぞれの資料における主題的な言葉である。動詞、形容詞は、意味の抽象度が高い言葉が挙げられる。連体詞「この」も、文章中でさまざまなものを示す言葉として多く用いられるもので、これらは基本的に文章の内容に関係なく、どのような場面にも用いられるものと見ることができることから、高頻度語には文章の骨組みとなる言葉が属しているといえる。

ランクの間であるランクⅢでは、形容詞に色彩語彙、感動詞には感謝などを表すもの、また、コミュニケーションに用いられるものが挙げられる。動詞も、具体的な意味を表すものが見られるようになる。名詞、代名詞には、具体的な概念や事柄を示す語が多く、動詞などの用言や、副詞、連体詞、接続詞、感動詞なども具体的な様子や関係を表す語が多い。低頻度語になるほど、そのような性質の語が多数を占めるのである。

**語種ランクと語種** ここでは語彙と語種の面から考える。

表8 語種のランク

ランク	和語	漢語	外来語	混種語
I	100.0%	0%	0%	0%
II	81.0%	4.9%	6.0%	8.1%
III	83.2%	9.2%	5.1%	2.5%
IV	75.0%	12.0%	7.7%	5.3%
V	74.2%	12.8%	5.6%	7.4%

表8から、ランクⅠはすべて和語で占められていることがわかる。ランクⅡ～Ⅴでは、漢語・外来語・混種語が見られるようになるが合わせても、20%～25%程度である。全体をとおして和語が多数を占める。日本語の語彙を語種の面から分析すると通常は異なり語数で漢語の比率が高くなり、述べ語数では和語の比率が高くなる傾向が知られており<sup>注14)</sup>、表8のように極端に和語に偏ることはない。

このように異なり語数で和語の比率が高いことは、幼児用絵本の特性と考えられるものである。

### 意味分類による分析

ここでは、「ぐりとぐら」シリーズの語彙を意味分類から考える。国立国語研究所（2004）『分類語彙表 増補改訂版』によって自立語を分類<sup>注15)</sup>すると表9のようである。

『分類語彙表』による分析では、抽象的関係の語が多く含まれることがわかる。以下の「 」内の語は使用例が多いものである。

表9 分類語彙表による分類

部門	語数	割合
1. 抽象的關係	357 語	40.5%
2. 人間活動の主体	27 語	3.1%
3. 人間の活動—精神および行為	239 語	27.1%
4. 生産物および用具	122 語	13.8%
5. 自然物および自然現象	137 語	15.5%
計	882 語	100.0%

1. 抽象的關係の言葉の使用が多く、「開ける 行く 来る 入る 飛ぶ 時 二匹 いっぱい 居る」など、時間・空間・作用・量・存在に分類される言葉が含まれる。

2. 人間活動の主体には、「お母さん 友達 僕 僕ら 私 世界」など、人間・社会に関わる語が見られる。語数は少ないが、これらの代名詞は子どもたちが自分と他者の区別するための語といえるものである。また2には、「ぎっく すみれちゃん<sup>注16)</sup> ぐら ぐり くるりくら」などの登場キャラクターの名前が含まれる。

3. 人間の活動—精神および行為には、「ありがとう 朝ご飯 歌う 大掃除 踊る 食べる 持つ わくわく」など生活に関する言葉が含まれる。日常生活における基本的な言葉と言えるものである。

4. 生産物および用具には、「お砂糖 カステラ 牛乳 ケーキ パン」など、子どもたちが日常口にする食品に関する言葉と、「襟巻き エプロン お鍋 毛糸 籠 雑巾 箒 帽子 リュック サック」など一般家庭にある道具や、衣服に関する言葉が含まれる。

5. 自然物および自然現象には、「青い 赤い お月様 白い そよそよ 光 ふわふわ 雪」など自然に分類される言葉、「枝 カボチャ 木 草 人参」など一般的に身の回りにあるものを表す言葉や、「足 頭 お腹 背中 手 目」など身体に関する言葉が含まれる。

言葉の意味カテゴリーの分析からも、子どもたちの身近にある言葉が多く用いられていることがわかる。これらは子どもたちの生活の場において必要かつ、身近な言葉である。

なお、表9の分析については下位も分類しているが、更に調査・考察が必要であるため課題とする。

### まとめと課題

子どもの言葉の育みを考えるために、子どもたちが最初に出会う本である絵本の中に見られる言葉を、「ぐりとぐら」シリーズをテキストとして数量的、意味的にデータ化し考察した。特に語彙の分野では、一般的な文章における語種の比率とかなり異なることがわかった。これは語彙ランクの

語種との関係においても同様である。

意味の分析においても、使用されている言葉は日常的な子どもたちの身近なものが多く見られた。本稿では「ぐりとぐら」シリーズのみの分析であるが、このような絵本と出会い、読み聞かされることで子どもたちは言葉を育み、学んだ言葉を用いてコミュニケーションにつなげていると考えられる。

これらの研究を進めるために、さらに多くの絵本のデータ分析が必要である。また、実際に、どのような単位で子どもたちは言葉を学ぶのかを考えると、やや細かい語彙分析ではなくもう少し長い単位での分析が必要であると考えられる。以後の課題としたい。

### 参考文献

- 小野正弘. (2007). *日本語オノマトペ辞典*: 小学館.
- 小野正弘 (主幹). (2024). *現代新国語辞典 第7版*: 三省堂.
- 国立国語研究所. (2004). *分類語彙表 増補改訂版*: 大日本図書.
- 平 博順・藤田早苗・小林哲生 (2012). *絵本テキストにおける高頻度語彙の分析*: 情報処理学会関西支部 支部大会 講演論文集.
- 田中牧朗・橋本行洋・小木曾智信編. (2021). *コーパスによる日本語史研究*: ひつじ書房.
- 中川素子・吉田新一・石井光恵 他 (編集). (2011). *絵本の事典*: 朝倉書店.
- 中川李枝子・大村百合子. (1963). *ぐりとぐら*: 福音館書店. (『母の友』1963.3所収)
- 中川李枝子・山脇百合子. (1966). *ぐりとぐらとおきやくさま*: 福音館書店.
- 中川李枝子・山脇百合子. (1976). *ぐりとぐらのかいすいよく*: 福音館書店.
- 中川李枝子・山脇百合子. (1979). *ぐりとぐらのえんそく*: 福音館書店.
- 中川李枝子・山脇百合子. (1987). *ぐりとぐらとくるりくら*: 福音館書店.
- 中川李枝子・山脇百合子. (2000). *ぐりとぐらとすみれちゃん*: 福音館書店.
- 中川李枝子・山脇百合子. (2002). *ぐりとぐらのおおそうじ*: 福音館書店.
- 中川李枝子. (2013). *月刊エモ4月号*: 白泉社.
- 中曾根仁・川又瑠璃子. (1994). *絵本の語彙*: 国立国語研究所.
- 日本国語大辞典編集委員会. (2000~2002). *日本国語大辞典 第2版*: 小学館.
- 藤田早苗・小林哲生・奥村祐子 他. (2017). *幼児の語彙獲得と絵本コーパスの関係を探る*: 言語処理学会 第23回年次大会 発表論文集.

### 注 釈

注1) 浅木尚実 (編). (2015). *絵本から学ぶ子どもの文化*: (pp98-112) 同文書院.

大久保 愛. (1993). *乳幼児のことばの世界* (p126) 大月書店.

中川素子・吉田新一・石井光恵 他 (編集). (2011). *絵本の事典* (pp510-519) 朝倉書店など.

- 注2) 日本では2000年の「子ども読書年」を契機に、翌2001年にスタートした。中川素子・吉田新一・石井光恵 他 (編集). (2011). *絵本の事典* : (p185) 朝倉書店.
- 注3) 平林一利. (2023). 『ぐりとぐら』に見られる言葉 絵本の語彙研究のための一考察. 豊岡短期大学論集 第19号, 171-179.
- 注4) 他に, 中川李枝子 (作)・大村百合子 (絵). (2009). *ぐりとぐらのしりとりのうた* : 福音館書店なども出版されているが, 一般的な物語形式とはことなるため対象としない.
- 注5) 「Web茶まめ」は, 「大学共同機関法人 人間文化研究機構 国立国語研究所」が開発した形態素分析ツールである. 日本語の語彙分析ツールとして, その信頼性から研究・教育に多く利用されているため, 本稿でも基礎解析に利用した.
- 注6) 実際には, 地の文, 発話文, 歌などに分けて分析しているが, 今回は一つにまとめている.
- 注7) 示したものはすべてではなく一部である. また, 語種についても, 複合させたばあい修正をしている.  
例 お料理 (和語 + 漢語) → 混種語
- 注8) 計量国語学会編. (2020). *計量国語学辞典 新装版* : (pp97-99) 朝倉書店.
- 注9) 注釈8と同.
- 注10) 計量国語学会編. (2020). *計量国語学辞典 新装版* : (pp82-85) 朝倉書店.
- 注11) 田中牧朗. (2016). *日本語研究法【近代語編】* : あおば言葉の会編, 『浮雲』の量的構造 (pp53-64) おうふう.
- 注12) ランクの分け方は便宜的なもので, 用例数の増減により区分も変化する.
- 注13) ランクが下がるにしたがって, 語彙数が増加するため, ランクⅢは使用度の高いものを示す.
- 注14) 注釈8と同.
- 注15) 『分類語彙表』に示されていない語彙に関しては, 分類項目一覧を参考にして分類した. また『分類語彙表』の中項目での分類をおこなった.
- 注16) 「ぐりとぐら」シリーズで実際の名前が使われるのはこの一語のみである. これは, 4歳で亡くなった女の子が『ぐりとぐら』を楽しんだことを作者が知り, そのイメージから執筆されたものとされる. 月刊モエ4月号 : (2013.3). 白泉社 p17による.