



話し言葉10種の言語データのコーパス調査をふまえた日本語学習者のためのABAB型基本口語オノマトペの選定

廉, 沢奇

(Citation)

統計数理研究所共同研究リポート, 456:117-129

(Issue Date)

2022-03-03

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCD01)

<https://doi.org/10.24546/81013073>

(URL)

<https://hdl.handle.net/20.500.14094/81013073>



話し言葉 10 種の言語データのコーパス調査をふまえた

日本語学習者のための ABAB 型基本口語オノマトペの選定

廉 沢奇(神戸大学大学院研究生)

Lzq980628@outlook.com

Selection of Basic Colloquial Onomatopoeia of ABAB-Type for Learners of Japanese:

A Corpus Study Based on 10 Kinds of Spoken Language Data
LIAN, Zeqi (Graduate School of Kobe University, Research Student)

概要

本研究は、日本語会話において頻出する基本的な ABAB 型オノマトペ(例:どンドン, そろそろ)の抽出を目指したものである。基本オノマトペの選定方法としては、「日本語日常会話コーパス」(CEJC)モニター公開版を用い、性別(女・男)・年代(子供・少年・中年・老年)・親疎関係(家族・親友・同僚・仕事)を区分した 10 種類の発話データごとに頻度調査を行い、全体での出現頻度にレンジ数を加味することで基本度を決定した。これより、227 語のオノマトペが抽出された。これらは、先行研究で示された書き言葉データに基づく基本オノマトペとは異なり、日本語の口語の特徴を反映したものとなっている。

キーワード

オノマトペ, 会話, 基本語彙, コーパス調査

1. はじめに

日本語の語彙体系において、オノマトペはきわめて重要な位置を占めており、小野(2007)によると、その数は 4000 を超えるという。また、新しいオノマトペも日々作り出されている。

一方、日本語教育の観点から見た場合、オノマトペは、語彙指導の直接の対象とならないことも多く、その正確な習得は決して容易ではない。下記の例を見てみよう。これらの用例は、「多言語母語の日本語学習者横断コーパス」に含まれる学習者の用例である。(1)はベトナム人(日本国内で生活している自然環境学習者)、(2)と(3)は中国人の用例である。なお、(1)は筆者が誤用と認定した用例であり、(2)と(3)の誤用の特定や修正例の提示はコーパス開発者の側でなされたものである。

(1)あとー, あと, どんどんどんどん, あー, 慣れました。(JJE03-I)

(2) 日本語一っつか、え一っつびぐり(×)した、わらしは高校時代は、んー日本語を、日本語を勉強していません、です。(CCH31-I)

(3) 今は学校の授業は、たぐさんたくさんがありますーから、それから、そのその(×) 春休み、春休みに、なるから(JJE50-RP1)

(1)はオノマトペの語形としては間違っていないが、この文脈(日本料理に慣れる過程について)であれば「だんだん」のほうが自然であると考え。また、(2)と(3)は「びっくり」と「そろそろ」というオノマトペの表記の誤りである。これらの用例は、多くの学習者にとって、オノマトペの形態面と意味面の理解が十分ではないことを示す。このように、オノマトペの習得が困難である背景には 2 つの原因が想定される。1 点目は、日本語教育において、オノマトペの指導が十分に行われていないことである。下記は中国の大学の日本語学科の一部で使われている『新経典日本語』の第 11 課の文法項目(「ので」)の解説欄に追加で掲載された用例である。

(4)もう時間なので、そろそろ帰ります。

上記はこの教科書における ABAB 型オノマトペの初出例であるが、ここには「差不多該」(ある時期・状態になりつつあるさま)という中国語での意味が添えられているだけで、オノマトペそのものの解説やその形態的特徴に関する説明は一切行われていない。中国に限らず、海外の日本語教育は教科書の単元に即して進むことが多いため、このように、各単元にたまたま出現したオノマトペに触れることはあっても、それらを体系的に学ぶ機会ほとんどない。

2 点目は、従来の日本語教育で紹介されるオノマトペは、どちらかと言えばフォーマルなものや書き言葉的なものが多く、自然な日本語会話で使えるオノマトペが十分に紹介されていないことである。このことは、オノマトペがぐだけた会話で多く使われるという日本語の実態をふまえると、問題だと言える。

以上の問題点をふまえると、現行の教科書を補完するものとして、現代日本語の話し言葉(会話)に焦点を当てた「基本口語オノマトペリスト」の作成がまずもって必要となろう。オノマトペには様々な形態があるが、本研究では、最も典型的な 4 モーラの ABAB 型オノマトペ(例:どんどん・がらがらん・ばりばり)に対象を絞った上で、基本口語オノマトペの選定を行う。なお、本研究は基準を満たす語形を網羅的に対象とするもので、意味の面でのオノマトペ性の判断(擬音語・擬態語的な用法に限定されるか、それに加え、固有の語彙的意味を持つかどうか)について、今回は行わないことにする。

2. 先行研究

小嶋(1972)によれば、「オノマトペ」はギリシア語の「*onomatopoiia*」を語源としており、「音による命名」ないし「音それ自身が名になる」の意を持つ。ただ、この定義に当てはまる語は多く、オノマトペに関する先行研究の多くが、形態や意味を手掛かりとしたオノマトペの分類や、重要オノマトペ

の抽出を目的としたものとなっている。以下では主要研究を概観する。

すでに述べたように、オノマトペの数は 4,000 を超えるとされる。こうしたオノマトペを議論する際には、それらをいくつかのグループに分類して分析することが広く行われてきた。

内容的分類に関して、金田一(1978)はオノマトペをその意味的な性質に基づいて分類した。オノマトペはまず、擬音語(外界の音を写した言葉「ザーザー」、「コツコツ」)と擬態語(音を立てないものを、音によって象徴的に表す言葉「キラキラ」、「サラサラ」)に分けられる。さらに、擬音語の中には擬声語(動物の鳴き声をまねして写した言葉「ワンワン」、「コロコロ」)もあり、擬態語の中には擬情語(人間の心の状態を表す言葉「イライラ」)と擬容語(生物の状態を表す言葉「きよろきよろ」)があるとされる。

また、オノマトペの音韻形態に基づく分類もある。まず、オノマトペ固有の音韻形態特徴に関して、Waida(1984)は「促音」「撥音」「り」「反復」をオノマトペの特徴として挙げ、それらを「オノマトペ標識」と呼んだ。田守(1999)は「母音の長音化」を一つのオノマトペ標識に加えた。また、浜野(2014)は/p/で始まるオノマトペが最も多く、逆に/r/で始まるオノマトペが最も少ないということを、日本語オノマトペの音韻特徴として挙げた。こうした研究をふまえ、日向・笹目(1999)はオノマトペ 1647 語をその語形から 49 種類に分類している。その結果、全体の 4 分の 1 を占め、最も多いのが、「わくわく」「がたがた」などの「ABAB」型で、以下、「どさっ」「にこっ」などの「ABっ」型、「ぐるり」「きらり」などの「AB り」型、「すっきり」「しっかり」などの「A っ B り」型、「がちゃん」「ぼかん」などの「AB ん」型が続くことが分かった。

また、オノマトペの語長や品詞性に着目した分類研究もある。内田(2015)は『日本語オノマトペ辞典』の「意味分類別索引」に掲載されたオノマトペ(1751 語)を抽出し、9 万語の『現代日本語書き言葉均衡コーパス』(BCCWJ)コアデータで頻度調査を行った。その結果、全体の 70%が 4 文字で構成されていること、また、全体の 88.8%が副詞であることが示された。

重要オノマトペの抽出について、三上(2007)は、辞書の収録語数 32600 語のうち、オノマトペは 791 語で全体の約 2%を占めることを明らかにした。このことを根拠として、全体で 4000 語を学ぶ日本語中級学習者であれば、その 2%、つまり 80 語程度のオノマトペを学ぶべきだと考えた。この発想をもとに、三上は 8 種の基本語彙リストを調査し、3 種以上に含まれていた 87 語のオノマトペを抽出し、内省で取捨選択を行うことで、教育用オノマトペとして 70 語を選定した。

獅々見(2016)は、4 冊のオノマトペ辞典を調べ、2 冊以上のオノマトペ辞典に出現したオノマトペ 273 語を抽出した。その後、日本語母語話者を対象とするアンケートで個々の語の親密度(見聞き、使用)を調査し、親密度の低い 21 語を除外した。最後に「BTS による多言語話し言葉コーパス」で頻度調査を行った。こうして得られた出現頻度データ、親密度(見聞き)データ、親密度(使用)データを主成分分析で合成し、全体を得られた主成分得点順に並べかえ、オノマトペの順位を決定した。こうして、順位付きの重要オノマトペ 252 語を提案している。

張(2020)は 6 種のオノマトペ辞典を調べ、最低 1 冊以上に含まれるオノマトペ(4189 語)と 2 冊以上に含まれるオノマトペ(1740 語)を抽出した。その後、BCCWJ で 4189 語の使用頻度を調査し、頻度 1 のものを排除したところ、4189 語は 2006 語になった。この 2006 語と 1740 語を比

較したところ、共通するのは 1463 語となった。最終的に、出現頻度に辞書における収録率をかけて調整した頻度で全体を並び替え、1463 種の重要オノマトペとした。

上記の研究で示された重要オノマトペのうち、本研究で扱う ABAB 型に合致するもの(一部を表示)は、下表のとおりである。

表 1 先行研究で選定された ABAB 型オノマトペの一部

三上リスト(上位 30 語)				
いらいら	うろうろ	からから	がたがた	がやがや
がんがん	きらきら	ぎりぎり	くるくる	ぐるぐる
げらげら	ごろごろ	さらさら	ざあざあ	じろじろ
すらすら	そろそろ	ぞろぞろ	どきどき	どんどん
にこにこ	のろのろ	はらはら	ばたばた	ばらばら
ふらふら	ぶらぶら	ぶるぶる	ぺこぺこ	ぺらぺら
獅々見リスト(上位 30 語)				
そろそろ	ぎりぎり	どんどん	にこにこ	だらだら
いらいら	どきどき	ぼろぼろ	わくわく	ごしごし
ばらばら	ぼさぼさ	うろうろ	すらすら	どろどろ
ぐちゃぐちゃ	めちゃめちゃ	ぺらぺら	べたべた	ばりばり
ぎゅうぎゅう	ぶらぶら	ぶかぶか	べらべら	ぼちぼち
ぐるぐる	ふわふわ	ふらふら	きょろきょろ	うとうと
張リスト(上位 30 語)				
どんどん	そろそろ	いらいら	ぎりぎり	ばらばら
どきどき	だんだん	にこにこ	きらきら	わくわく
ぐるぐる	ぼろぼろ	ごろごろ	ふらふら	さらさら
うろうろ	くるくる	ぶらぶら	ばたばた	ぶつぶつ
だらだら	ふわふわ	がんがん	にやにや	ばりばり
ぴかぴか	たんたん	こつこつ	がたがた	ずるずる

上記で見たように、重要オノマトペ選定に関する先行研究は、使用するデータの選定、重要語抽出の手法、最終的に得られたリストの各点においてそれぞれ価値ある知見を示している。しかしながら、全体として書き言葉に偏った研究が多く、話し言葉を調査したものも、発話者の属性の違いにまでふみこんだ頻度調査はなされていない。

そこで本研究は、会話コーパスを調査資料とし、発話者属性別に頻度調査を行うことで、新しい観点から、重要な ABAB 型口語オノマトペの選定を目指す。

3. リサーチデザインと手法

3.1 研究目的と RQ

すでに述べたように、本論文の目的は、会話コーパスを分析し、発話者属性によるオノマトペ使用量の差を確認した上で、多様な属性の話者が共通して発話で使用するオノマトペを抽出し、今

後の日本語学や日本語教育で活用しうる重要会話オノマトペリストの作成を行うことである。なお、議論の対象は 4 モーラの ABAB 型オノマトペ(「どんどん」「きょろきょろ」など)に限定する。この目的に沿って、以下の 3 つの研究設問を設定した。

- RQ 1 性別・年齢・親疎関係により、オノマトペ使用量(延べ語数・異なり語数)は変化するか？
(話者属性別使用量)
- RQ 2 幅広い性別・年齢・親疎関係の発話者が共通して使用する「重要会話オノマトペ」にはどのようなものがあるか？(重要オノマトペ抽出)
- RQ 3 「重要会話オノマトペ」は、先行研究で選定された教育的オノマトペ(109 語)および書き言葉系オノマトペ(563 語)とどのように異なっているか？(重要オノマトペ特性)

3.2 データ

本研究で使用したコーパスは『日本語日常会話コーパス』(CEJC)である。CEJC はさまざまな場面における自然な会話を収集することを目的としたコーパスである。性別・年齢などのバランスを考慮して抽出された 40 名の調査協力者を集め、機材機器を約 3 か月間貸与し、協力者の日常生活で自然に生じる会話を記録させた。発話者数は全体で約 250 名となる。今回使用したデータはモニター版の 50 時間分の会話である。

CEJC の会話データは、会話のメタ情報(「形式」「場面」と、話者の属性情報(「性別」「年代」「会話相手との関係」))によってジャンルに区分されている。本研究は、「話者属性」情報を会話内容の分類基準とみなし、CEJC のデータを「女」「男」「子供」「少年」「中年」「老年」「家族」「友人」「同僚」「仕事」という 10 種の話者属性にわけて調査する。

本研究で分析する 10 種の基本データを下記に示す。なお、個々の総語数については、『日本語日常会話コーパス』モニター公開版の語彙(2020)による。

表 2 10 種発話者属性の基本データ

分類観点	発話者属性	延べ語数	異なり語数
性別	女	349,144	10,618
	男	260,183	9,841
年齢	子供	6,545	1,150
	少年	107,826	7,701
	中年	373,184	23,940
	老年	117,956	10,363
親疎	家族	225,192	-
	友人	278,418	-
	同僚	59,600	-
	仕事	46,117	-

3.3 手法

まず、RQ1 では、CEJC からすべての 4 モーラ ABAB 型オノマトペを抽出する。具体的には、Python を用い、元データから 2 文字の繰り返し(「どんどん」など)と 3 文字の繰り返しを抽出し、後者については 4 モーラ以外のもの(「べろんべろん」など)を削除し、4 モーラのもの(「きょろきょろ」など)を残す。これにより 227 語のオノマトペを得た。次に、これらの語を対象に、性別(女・男)、年齢(子供・少年・中年・老年)、親疎(家族・友人・同僚・仕事)の属性ごとにオノマトペの総頻度(延べ語数および異なり語数)を調査する。頻度の差は検定で確認を行うが、親疎間については総語数のデータが入手できなかったため、延べ語数についてのみ検定を行う。

次に、RQ2 では、CEJC において出現した 227 語のオノマトペについて、10 種の話者属性(女・男・子供・少年・中年・老年・家族・友人・同僚・仕事)ごとに使用頻度を調査し、それらを 100 万語あたりに調整した後に平均化する。その後、様々な発話環境で共通して高頻度に使用される語を選ぶため、平均頻度にレンジ比率($n/10$)をかけた値を重要度指標とみなし、指標の降順で全体を並べ替えて、重要度が 1.0 以上になるものを「重要会話オノマトペ」とする。以下は、3 語を例とした計算手順の実例である。

表 3 オノマトペの各ジャンル出現確率及び重要度の計算

	女	男	子供	少年	...	平均	レンジ	重要度
どんどん	100.2	261.3	458.3	120.5	...	215.9	10/ 10	$215.9 \times 1.0 = 215.9$
だんだん	71.6	92.2	0.0	27.8	...	80.4	9/ 10	$80.4 \times 0.9 = 72.36$
そろそろ	42.9	49.9	152.7	46.3	...	49.5	9/ 10	$49.5 \times 0.9 = 44.55$

RQ3 では、上記で抽出された重要会話オノマトペの妥当性と独自性を検証するため、先行研究で抽出されたオノマトペリストとの比較を行う。まず、三上(2007)、獅々見(2016)、張(2018)のリストに含まれる 4 モーラ ABAB 型オノマトペを抽出したところ、三上(2007)から 31 語、獅々見(2016)から 95 語、張(2018)から 563 語、全体で延べ 689 語が得られた。重複を削除すると、異なり語数は 564 語となる。これを、本研究で得られた重要会話オノマトペと比較し、本リストのみに含まれている語を抽出して内容的に検討する。

4. 結果と考察

4.1 RQ1 話者属性別使用量

性別・年齢・親疎の 3 観点に即して、話者属性別のオノマトペ全体使用量を調査したところ、以下の結果を得た。

図 1 性別のオノマトペ使用量(左:延べ語数, 右:異なり語数)

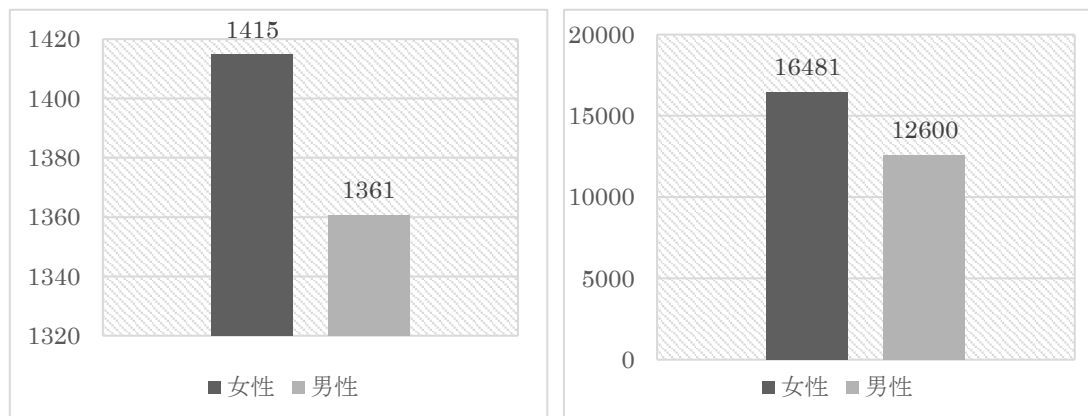


図 2 年齢別のオノマトペ使用量(左:延べ語数, 右:異なり語数)

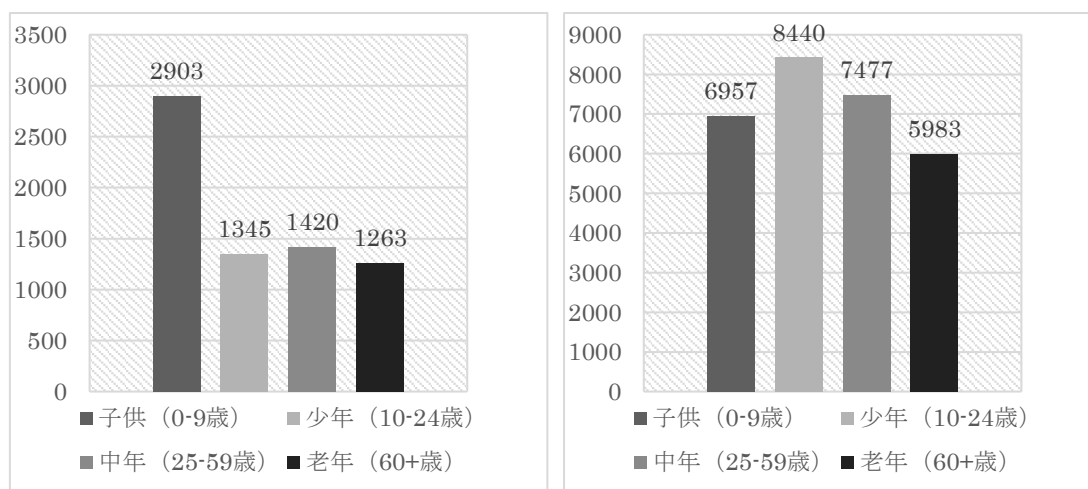
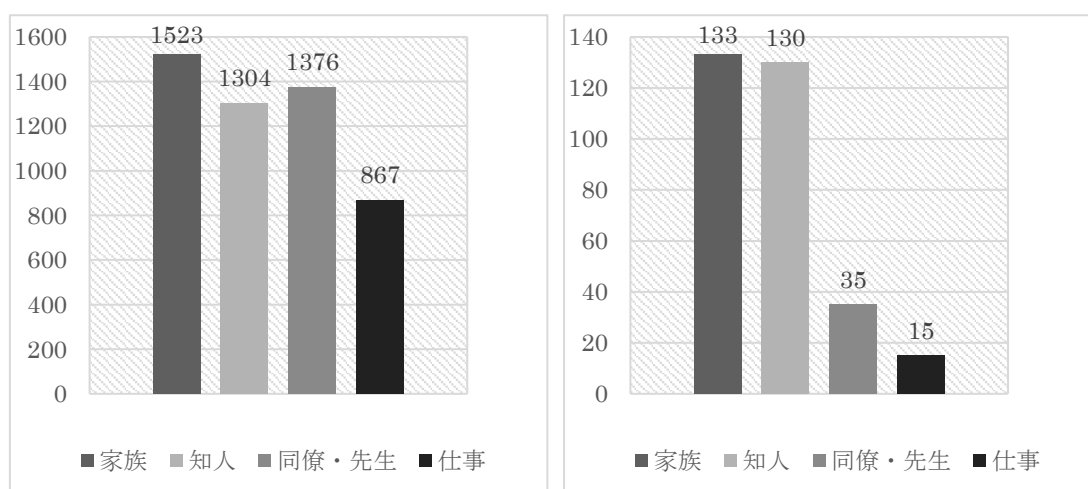


図 3 親疎別のオノマトペ使用量(左:延べ語数, 右:異なり語数)



まず、性別については、延べ語数・異なり語数とも女性がオノマトペをやや多く使用する傾向がみられたが、対数尤度比検定の結果、異なり語数のみ、差は有意であった(延べ語数: $G^2 = 0.32$, $p = 0.57$, 異なり語数: $G^2 = 5.38$, $p = 0.02$)。このことより、女性のほうが幅広いオノマトペを使用する可能性が示唆された。

最後に、親疎については、延べ語数、異なり語数ともに、家族間においてオノマトペを多用する傾向がみられた。分散分析を行ったところ、延べ語数において、年齢の主効果は有意となり($F(3,678) = 3.188, p = .033, \eta_p^2 = .014$)、事後検定(Holm)により、家族≒同僚≒知人>仕事の関係にあることが確認された。つまり、家族間において使用量が増えるのではなく、仕事関係の相手と話す場合に、他の場合より、オノマトペの使用量が全体として低下することが確認された。

4.2 RQ2 重要オノマトペ抽出

なお、上位 10 語は、「どんどん」「ぐるぐる」「だんだん」「そろそろ」「くるくる」「ふわふわ」「ぎりぎり」「しゅわしゅわ」「めちゃめちゃ」「しゅっしゅっ」であった。以下はそれらの用例である。(各々について 2 例を示す)

女性 50-54 歳)

(8b) 抜かなきゃじゃん, それたぶんだんだんだんだん前歯がクラッシュしてくるよ。(T009_014a

男性 20-24 歳)

(9a) 毎年冬ス寒さ厳しくなるとそろそろかなって。(S002_004 男性 45-49 歳)

(9b) いんですかじゃあ, そろそろ終わりにします。(C002_015 女性 55-59 歳)

(10a) 一二三て回してさくるくるくるくるくるって回して。(T003_012 男性 5-9 歳)

(10b) キュウリのこのくるくるってなってる端っこのこと。(T013_014a 女性 20-24 歳)

(11a) いただきます。ふわっふわ, うん, ふわふわうまい。(T006_004 男性 25-29 歳)

(11b) 泡でふわふわとかって落ちないって。(C002_015 女性 55-59 歳)

(12a) ぎりぎりの生活を送ってる。(K003_005 女性 20-24 歳)

(12b) なんかぎりぎりわかるんだよな。(T005_008 男性 35-39 歳)

(13a) これ回すと入んの, すごいえやすごいうおー, しゅわしゅわしゅわしゅわ。(T005_007 女性 65-69 歳)

(13b) まあ炭酸ですからね, そうっす, だいじょぶっす, あ, しゅわしゅわする。(T002_020 男性 40-44 歳)

(14a) 来年使ってくれ, めちやめちや使います。(T009_021b 女性 20-24 歳)

(14b) めちやめちや種類があるんでしょ。(T008_006 男性 30-34 歳)

(15a) 一枚一枚区切れんの, これせつせとしゅっしゅってやってくれればいいのに。
(K003_017 女性 20-24 歳)

(15b) えっと紙製のしゅっしゅっ紙切り切れる。(T003_015 男性 5-9 歳)

上記の用例はいずれも日本語としてごく自然なもので, 発話者の性別・年齢も幅広い。この点をふまえると, 本リストは日本語会話の諸相を反映した重要オノマトペをまとめたものとして, 今後の日本語会話教材開発の基盤資料になりうるだろう。

4.3 RQ3 重要オノマトペ特性

本研究で抽出された会話重要オノマトペ(全 109 語)を, 三上(2007), 獅々見(2016), 張(2018)の3種の先行リストに含まれる564種と比較した結果, 本リスト(全 109 語)のみに含まれる独自語が以下の18語であることがわかった。本リストのサイズは約100語と, 先行研究のリストに

比べて小規模なものだが、そのうち、約 2 割が、既存リストにない新しいオノマトペとなっていることになる。なお、下表では、本リストにおける順位帯ごとにわけて表示している。

表 4 本リストで独自に抽出された重要オノマトペ

ランク	該当オノマトペ
上位 20 位	しゅわしゅわ／そこそこ*／どぼどぼ
上位 50 位	とうとう*／つつい*／らくらく*
上位 80 位	ちゃーちゃー／ぶちぶち／わたわた／ぜいぜい／うじょうじょ
上位 81 位以降	ちょんちょん／わちゃわちゃ／しゃわしゃわ／かぷかぷ／くしゅくしゅ／ がっがっ／ぬりぬり*

この中には、「そこそこ」や「つつい」のように接尾辞や副詞として一定の語彙的意味を有するものや、「とうとう(到頭・滔々・等々)」、「らくらく(楽々)」、「ぬりぬり(塗り塗り)」のように、漢字表記が可能で、元の語の意味合いが残存しているものなど、オノマトペ性が相対的に低いものも含まれている。ただ、そうした語を除いたとしても、これまでのリストでは重視されていなかった新しい会話的オノマトペを一定数抽出することができたと考える。

もっとも、本リストで新しく選ばれたオノマトペは、決して、会話にしか使えない特殊な語というわけではない。ここでは、本リストの上位 80 位以内に入ったもののうち、擬音語・擬態語性が高いものとして、「しゅわしゅわ」(上記の(13a)と(13b)ですでに示した)「どぼどぼ」「うじょうじょ」「ぶちぶち」の 4 種を取り上げ、以下、用例を 2 つずつ示しておきたい。

(16a) そのままだどぼどぼ入れちゃって渡すんだけど。(T003_015 男性 5-9 歳)

(16b) こうどぼどぼどぼとウイスキーを注ぎ入れます。(日本語話し言葉コーパス S00F0476 女性 30-34 歳)

(17a) ジングルですから、ワニがうじょうじょいるところに、そのワニに食い付かれるんです。
(日本語話し言葉コーパス S07F0567 女性 65-69 歳)

(17b) なんかもう鳥がうじょうじょうじょうじょなんかね(T015_006 男性 50-54 歳)

(18a) 腰が痛くなったとか、膝が痛いとか、ぶちぶちぶちぶちこぼすわけ(T004_006 女性 60-64 歳)

(18b) 一二三、あ、そうだね、うん、ぶちぶちって出して(T015_006 女性 50-54 歳)

これらは、いずれも一般的な文脈でも使用可能なもので、日本語会話において汎用性を有する重要語と言えるだろう。こうした語がうまく抽出できていることが、本リストの価値であると考えられる。

5. まとめ

本研究は日本語日常会話コーパス CEJC を分析し、発話者属性によるオノマトペ使用量の差を確認した上で、多様な属性の話者が共通して発話で使用する重要会話オノマトペを抽出し、今後の日本語学や日本語教育で活用しうる「重要会話オノマトペ」リストの作成を行った。

RQ1(話者属性別使用量)では、発話者属性によるオノマトペ使用量の差を調査した。その結果、性別については、女性は男性より幅広いオノマトペを使用するが、総使用量は差がないことが分かった。年齢については、子供がオノマトペを全体として最も多用することが分かった。話し相手との親疎関係については、仕事関係の相手と話す場合にオノマトペの使用量が他より少なくなることが分かった。

RQ2(重要オノマトペ抽出)では、話者属性ごとのデータに基づいて重要度指標を計算し、全体を指標の降順で並べ替え、指標が 1.0 を超える 109 語の「重要会話オノマトペ」リストを作成した。

RQ3(重要オノマトペ特性)では、本研究で得られた「重要会話オノマトペ」を先行リストの掲載語と比較した結果、109 語中、約 2 割において、先行リストに含まれていない新しいオノマトペを抽出できていることが確認された。

以上で見たように、本研究は一定の成果を挙げたが、もちろん、課題も多い。もっとも重要な限界は、ベースにしたデータ量の制約である。この点については、今後、さらに多くの会話データを収集することで、リストの改善につなげていきたいと考える。

引用文献

- Waida, T. (1984). English and Japanese onomatopoeic structures. *A Journal of western Languages and Cultures*, 36, 55-79.
- 石川慎一郎・李在鎬・砂川有里子(2018)『新・日本語教育のためのコーパス調査入門』くろしお出版.
- 内田ゆず・高丸圭一・乙武北斗・木村泰知(2016)「BCCWJ におけるオノマトペの語義とコロケーションの分析」『人工知能学会全国大会論文集』1-4.
- 小野正弘(2007)『擬音語・擬態語 4500 日本語オノマトペ辞典』小学館.
- 金田一春彦(1978)『擬音語・擬態語辞典』角川書店.
- 獅々見真由香(2016)「日本語の会話におけるオノマトペの基本語彙選定」『日本語教育』165,73-88.
- 田守育啓・ローレンススクウラップ(1999)『オノマトペ:形態と意味』くろしお出版.
- 張晶鑫(2020)「現代日本語におけるオノマトペの用法解明と中国人日本語学習者のためのオノマトペ指導に対する提言—コーパス言語学の教育的応用の可能性をめぐる—」神戸大学国際文化学研究科博士論文.
- 浜野祥子(2014)『日本語のオノマトペ-音象徴と構造』くろしお出版.
- 日向茂男・笹目実(1998)「語形からみた擬音語・擬態語 2」『東京学芸大学紀要第 2 部門人文科学 50』189-209.

三上京子(2007)「日本語教育のための基本オノマトペの選定とその教材化」『ICU 日本語教育研究』49-63.

渡邊裕子(1997)「日本語教育におけるオノマトペの扱いについての一考察」『兵庫教育大学学校教育研究』9,23-31.

附表 日本語重要会話オノマトペリスト(全 109 語)

順位	オノマトペ	レンジ	重要度	順位	オノマトペ	レンジ	重要度
1	どんどん	10	215.9	2	ぐるぐる	8	83.1
3	だんだん	9	72.4	4	そろそろ	9	44.6
5	くるくる	7	36.9	6	ふわふわ	8	33.6
7	ぎりぎり	8	31.5	8	しゅわしゅわ	7	30.7
9	めちゃめちゃ	8	30.0	10	しゅっしゅっ	7	24.2
11	がんがん	8	21.8	12	そこそこ	9	16.2
13	ばりばり	8	10.1	14	どぼどぼ	3	9.7
15	ぼんぼん	7	9.1	16	ちょろちょろ	6	8.8
17	ぼんぼん	6	7.9	18	びしょびしょ	6	7.5
19	がたがた	7	7.5	20	ぼろぼろ	5	7.3
21	ばんばん	7	7.1	22	どきどき	6	7.0
23	ちよきちよき	6	6.7	24	ばらばら	7	5.4
25	ばたばた	6	5.2	26	いらいら	6	5.1
27	ぼちぼち	5	4.8	28	びよびよ	3	4.8
29	ちょこちょこ	7	4.7	30	つるつる	6	4.4
31	にやにや	6	4.2	32	べらべら	7	4.1
33	ごろごろ	7	4.1	34	とうとう	6	4.1
35	どろどろ	6	3.9	36	ちょいちょい	6	3.9
37	にこにこ	6	3.9	38	つつい	6	3.4
39	とろとろ	6	3.3	40	くりくり	5	3.2
41	もちもち	6	3.2	42	らんらん	5	3.2
43	べたべた	5	3.1	44	もこもこ	5	3.1
45	ぱさぱさ	5	2.9	46	がりがり	5	2.8
47	かたかた	5	2.6	48	すかすか	5	2.6

49	てかてか	5	2.5	50	らくらく	5	2.4
51	わくわく	5	2.3	52	ばんばん	5	2.3
53	ばしゃばしゃ	5	2.3	54	こりこり	5	2.2
55	しくしく	4	2.2	56	うろうろ	5	2.1
57	じろじろ	5	2.0	58	つんつん	5	2.0
59	よれよれ	6	1.9	60	ぶるぶる	4	1.9
61	ちゃーちゃー	5	1.9	62	ごちゃごちゃ	5	1.8
63	うはうは	4	1.8	64	ぶちぶち	4	1.8
65	わたわた	3	1.6	66	かちかち	4	1.6
67	ぜいぜい	6	1.6	68	ぼさぼさ	6	1.6
69	ぐしゃぐしゃ	5	1.6	70	ぱっぱっ	6	1.6
71	かちやかちや	3	1.5	72	くいくい	5	1.5
73	たんたん	3	1.5	74	とんとん	4	1.5
75	ぶーぶー	5	1.5	76	がちやがちや	5	1.5
77	けらけら	3	1.4	78	うじょうじょ	3	1.4
79	わいわい	5	1.4	80	ころころ	4	1.4
81	もやもや	4	1.4	82	ちよんちよん	5	1.3
83	どばどば	3	1.3	84	ねばねば	5	1.3
85	ぷりぷり	4	1.3	86	わちやわちや	5	1.3
87	ざわざわ	4	1.3	88	ぴりぴり	3	1.3
89	くねくね	5	1.3	90	もくもく	5	1.3
91	ヒーヒー	4	1.2	92	しゃかしやか	4	1.2
93	しゃわしゃわ	4	1.2	94	ぐちゃぐちゃ	4	1.2
95	まじまじ	4	1.2	96	ぎざぎざ	4	1.1
97	ふにやふにや	4	1.1	98	ばらばら	4	1.1
99	びろびろ	3	1.1	100	かふかふ	3	1.1
101	ごとごと	4	1.1	102	くしゅくしゅ	3	1.1
103	がっがっ	3	1.1	104	おどおど	3	1.1
105	ぬりぬり	3	1.1	106	こつこつ	3	1.0
107	はあはあ	5	1.0	108	かりかり	5	1.0
109	きらきら	5	1.0				