



書き言葉・話し言葉を含む72変種のコーパスデータの調査に基づく「重要漢語動名詞200」の抽出の試み

陳, 迪

(Citation)

統計数理研究所共同研究リポート, 456:21-41

(Issue Date)

2022-03-03

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/81013066>

(URL)

<https://hdl.handle.net/20.500.14094/81013066>



書き言葉・話し言葉を含む 72 変種のコーパスデータの調査に基づく

「重要漢語動名詞 200」の抽出の試み

陳 迪 (神戸大学大学院研究生)

chendi24410@gmail.com

Choosing 200 Important Sino-Japanese Verbal Nouns Based on the Frequency Analysis of 72 Kinds of Spoken and Written Corpus Subsets CHEN, Di (Graduate School of Kobe University, Research Student)

概要

教育的に重要な日本語漢語動名詞(例:努力・協力・勉強)を抽出する試みは先行研究で広く行われているが, その多くは書き言葉のデータを大きなかたまりとして分析したもので, 話し言葉を含む現代日本語の多様なジャンル差に十分配慮した研究は必ずしも多くなかった。本研究は, 「現代日本語書き言葉均衡コーパス」と「日本語日常会話コーパス」に含まれる言語データを 72 の変種に分類した上で, 変種ごとに漢語動名詞の出現状況を網羅的に調査し, 多様な現代日本語における漢語動名詞の使用実態を解明するとともに, 汎用的な初中級学習者向けの重要漢語動名詞の抽出を試みた。その結果, RQ1(変種別頻度)について, (1)変種間の頻度差は最大 1284 倍にのぼること, (2)漢語動名詞は大量の情報を効率的に伝達することを優先する変種で多く, 話し言葉や, 低学年向け教科書など, 双方向的に情報を伝達し, 読み手・聞き手間の共感醸成を重視する変種で少なくなることがわかった。RQ2(最頻出語の重複度)では, 変種別の高頻度漢語動名詞の重なり度合いは低く, 多くの漢語動名詞が特定の変種に従属することが明らかになった。RQ3(重要漢語動名詞の抽出)では, 「平均頻度×レンジ比率」という尺度に基づき, 重要漢語動名詞 200 語を新たに抽出した。最後に, RQ4(重要漢語動名詞の教育的妥当性)では, 本研究で得られたリストが初中級の学習者向けの教育用語彙リストとして妥当であるかを 2 段階で検証した結果, 本リストには, 従来のリストで欠落していた初級者向けの重要漢語動名詞が盛り込まれていることが確認された。

キーワード

漢語動名詞, 重要語, コーパス

1. はじめに

本稿で扱う漢語動名詞とは, 「努力」「勉強」のように, 名詞自体に動詞的な意味を有し, 「する」を付加することで動詞としても使用できるものである。漢語動名詞は, 古くから言語学的な関心を引き

付けてきた。たとえば、山田(1940)は、『国語の中に於ける漢語の研究』において、漢語動名詞を「外来語の本義が動詞なるものの転来せるもの」で、「国語にては一種の名詞として取り扱われる」ものと定義している。なお、漢語動名詞は、「する」を伴って動詞として使用される場合、「漢語サ変動詞」と呼ばれるが、本研究では、当該漢語の名詞的な使用と動詞的な使用をあわせて分析することから、以下、漢語動名詞という呼称で統一する。

漢語動名詞の振る舞いは、従来、言語学的関心から広く調査されてきたが、漢語動名詞は日本語教育の観点からも興味深い研究対象となりうる。というのも、漢字文化を有する中国人学習者を含め、多くの日本語学習者が漢語動名詞の習得に困難を覚えているためである。以下の例を見てみよう。これらは、学習者作文コーパス「なたね」(<https://hinoki-project.org/natane/>)に収録されている中国人学習者の作文実例である。なお、修正例はコーパス開発者が付与したものである。

- (1) (15135) 人口の情勢がいい国に進んだ経験を勉強します (→学びます)。
- (2) (18498) 西洋の文化を習いながら、片仮名で表示する (→表記する) 外来語がどんどん多くなっている。
- (3) (15049) 中国の茶文化はこのようにだんだん形成しました (→形成されました)。

これらの誤用は、豊富な漢字の知識を有する中国人学習者であっても、漢語動名詞の習得が容易ではないことを明確に示す。

では、なぜ、漢語動名詞の習得が困難なのであろうか。この点については、少なくとも3つの理由が考えられる。1点目は、日本語教材において漢語動名詞が十分に指導されていないことである。教科書に出てくる日本語漢語については、多くの場合、読み方・アクセント・品詞性・動詞タイプ・中国語訳などが記載されるが、漢語動名詞の名詞的・動詞的用法を体系的に説明することは少なく、紹介される関連用例も限られている。以下は『新日本語教程(初級1)』(人民教育出版社)の第6課の記載例である(pp.54-55)。

- (4) 法子：ごめんください。
園長：こんにちは。見学ですか。わたしはこの幼稚園の園長です。
※見学(けんがく)① [名・他動3] 参观, 参观学习

漢語動名詞の「見学」について、品詞と意味は示されているが、動名詞としての実際的な使い方までは触れられていない。

2点目は、中国人日本語学習者の学習目標とされる各種の日本語テストにおいて、漢語動名詞の知識を直接に問う問題が少ないことである。以下は『日本語能力試験公式問題集第二集』(日本語能力試験公式ウェブサイト <https://www.jlpt.jp/>)に掲載された問題実例である。

- (5) は N4, (6) は N3, (7) は N2 の「文字・語彙」セクションの問題である。

(5) さとうさんが けがをしたと 聞いて、みんな () しました。

- 1 しんぱい 2 けいけん 3 しつれい 4 おじぎ (p.4)

(6) 沸騰

- 1 今日は朝からどんどん暑くなり、昼には気温が沸騰した。
2 鍋のお湯が沸騰したら、とうふを入れて火を少し弱くしてください。
3 昼ごろから具合が悪くなり、夕方熱が沸騰したので病院へ行った。
4 このストーブは沸騰するのが早いので、すぐに部屋が暖かくなる。(p.5)

(7) わが社の商品はここでせいぞうされている。

- 1 製増 2 制増 3 制造 4 製造 (p.2)

上記の (5) と (6) は漢語動名詞の意味を問う問題、(7) は漢語の読み方と同音異義語の使い分けを問う問題であるが、いずれもそこに含まれる漢語動名詞の振る舞いを直接に問うことはなされていない。

3点目は、学習者が参照する日本語辞書において漢語動名詞の情報が十分に提供されていないということである。そもそも、一般の辞書には頻度の情報やジャンルの情報が掲載されていないため、学習者は、辞書を見ても、当該の漢語動名詞が現代日本語の各ジャンルにおいてどの程度使われているかを確認することはできない。以下は『新明解国語辞典(第七版)』における「規定」の解説である。アクセント関係の記述は省略している。

(8) きてい【規定】

[一する] (他サ)

<なにヲ (なんだト) 一する>

⊖判断のよりどころとするために、あいまいな内容のある決まった形に定めること。

「前項に一する通り...／一量」

⊖行動の是非を判断するよりどころとなるべき個々の規則や条文 (を定めること)。

「一を設ける／一に従う」

日本語の「規定」は書き言葉的な語であり、類義の「約束」や「ルール」と異なり、一般のくだけた話し言葉で使われることは少ないと思われるが、そうした情報は辞書にも掲載されていない。中国語では「規定」が話し言葉でも使えるため、中国人学習者は、語のフォーマリティを意識せず、日本語の話し言葉で一般語として「規定」を使ってしまう危険もある。

以上で、漢語動名詞の習得の困難性の原因を考えてきた。こうした背景をふまえると、今

後の日本語教育においては、漢語動名詞の習得をいわゆる偶発的語彙学習 (incidental vocabulary learning) だけにゆだねず、重要な漢語動名詞をリストアップした上で、他の文法項目や和語・外来語などと同じく、独立した指導の対象として位置付けることが不可欠であろう。

2. 先行研究

本研究で扱う漢語動名詞は、サ変動詞の語幹部分に相当する。本節では、はじめに、漢語動名詞の呼称に関わる実態を概観した後、本稿の趣旨に沿って、重要漢語動名詞の抽出に関わる主要な先行研究を紹介する。

まず、漢語動名詞については、主要な研究・文法体系・形態素解析辞書において、以下のような呼称が使用されてきた。

表1 漢語動名詞の呼称

研究・文法体系・形態素解析辞書	呼称
平尾 (1990), 影山 (1993), 小林 (2004) 等	動名詞 (Verbal Noun)
間淵 (2016)	サ変可能名詞
学校文法, 日本語教育文法	名詞
IPA 辞書	名詞・サ変接続
UniDic 辞書	名詞・普通名詞・サ変可能

前述の山田(1940)のほか、平尾(1990)も、サ変動詞の語幹部分を動詞的な名詞とみなし、それを「動名詞」(Verbal Noun, VN)と呼んでいる。影山(1993)は、「動名詞」を日本語の主要な語彙範疇であるとみなしている。小林(2004)もその立場をふまえ、「動名詞」という名称を用い、その名詞的特徴と動詞的特徴をまとめている。

一方、間淵(2016)は、サ変動詞用法を持つ2字漢語名詞を「サ変可能名詞」と呼んでいる。

学校文法や教育文法の品詞体系では、「研究」のように、名詞として使えるとともに「する」を付けると動詞としても使えるものを名詞として扱っている。その語に「する」を付けた形は、学校文法では「サ行変格活用動詞」、教育文法では「Ⅲ類動詞」と呼ばれる。

情報処理振興事業協会 (IPA) が定める品詞体系 (THiMCO97) では、漢語や外来語、和語名詞などの後ろに「する」「できる」「なさる」「くださる」などを後接させられるものを「名詞・サ変接続」と呼ぶ (『ipadic version 2.7.0 ユーザーズマニュアル』, p.19)。

また、「国立国語研究所短単位自動解析用辞書 (UniDic)」は IPA 品詞体系を参考して、学校文法の品詞・活用型・活用形を細分化して階層的な品詞体系を設計した。UniDic では、普通名詞のうち、「研究」のように「する」「できる」などが直後に続き、動詞として使われるものを「名詞・普通名詞・サ変可能」と呼んでいる (『UniDic version 1.3.9 ユーザーズマニュアル』, p.15)。

以上で概観したように、当該品詞については様々な呼び方・とらえ方があるわけだが、本稿では、名詞と動詞の両方の特徴を有する漢語サ変動詞の語幹を指して「漢語動名詞（漢語VN）」という呼称を使用する。

続いて、重要漢語動名詞の抽出に関する研究を見ていく。まず、李（2013）は「現代日本語書き言葉均衡コーパス」（BCCWJ）を用いて高頻度・汎用的漢語サ変動詞を抽出した。調査した言語変種は BCCWJ の書籍、雑誌、新聞、ブログ、白書、知恵袋の 6 種であるが、書籍については日本十進分類法（NDC）による 10 ジャンルをそれぞれ独立した変種とみなし、あわせて 15 変種のデータに対して調査を行った。調査対象語として、2 文字からなる漢語サ変動詞に絞った。具体的には、まず、15 変種ごとに高頻度上位 50 語を抽出し、重複を除いた 279 語を決めた。次に、15 変種のうち 1 変種しか上位 50 語に入っていないものを除き、277 語を対象として変種別の頻度を調査し、漢語サ変動詞の汎用性を確認した。最後に、主成分分析を用いて 15 変種の頻度を合成し、高頻度・汎用的な漢語サ変動詞 93 語を特定した。

張（2014）は「読売新聞コーパス」の 2000 年に発行されたデータを用い、タグを除き、「茶筌」で形態素解析し、得られた結果を元データ（「第 1 次データ」）として二字漢語動詞を抽出した。ここで言う漢語動詞とは、「漢語動名詞＋する」を指すものである。具体的な抽出手法は、(1) perl で作成したプログラムを利用して「第 1 次データ」から全てのサ変動詞を抽出し、後ろに「する、させる、される、できる」及びその活用形が来る語を「第 2 次データ」とし、(2) 「第 2 次データ」で形式の同じものを一つにしてその出現頻度を集計し、得られた異なり語数 36840 語を「第 3 次データ」とし、(3) 「第 3 次データ」における外来語のサ変動詞と和語のサ変動詞を除き、漢字が 2 字であるものを抽出し、異なり語数で 4433 語を「第 4 次データ」とし、(4) 「第 4 次データ」で音読みにならないものを手作業で除いた。その結果、4383 語の二字漢語動詞を抽出した。さらに、これらの漢語動詞に対して、語構成パターン（AV 型・VN 型・VV 型・MV 型・接辞型・その他）、自他パターン（例内・例外）、語彙レベル（旧日本語能力試験語彙リストにおけるレベル）の情報を付加した。

間淵（2016）は、現代語コーパスとして BCCWJ の出版サブコーパスと、近代語コーパスとして「明六雑誌コーパス」「国民之友コーパス」「太陽コーパス」「近代女性雑誌コーパス」を用い、近代・準近代・現代・準現代・通時的二字漢語サ変動詞を抽出した。氏は調査対象語選定の抽出域を「二字漢語全体」とした上で、キー条件について、「語種」を「漢語」、「語彙素」の文字数を 2 文字に設定し、後文脈条件について、「語彙素」を「為る」「出来る」「致す」「奉る」「仕る」「遊ばす」「下さる」「なさる」「申す」に設定した。これにより、いずれかのコーパスにおいてサ変動詞用法を持つと思われる二字漢語サ変可能名詞として約 12000 語を抽出した。それらの語の品詞性を一定の条件によって認定し、サ変可能以外の語を削除し、近代・現代コーパスでのサ変動詞用法の有無によって整理類別した。その結果、近代・現代に共通してサ変動詞用法が見られる通時的サ変は 2043 語であった。近代ではサ変動詞用法がほぼなかったものの、現代になって新たにその用法を獲得した準現代・現代サ

変は 82 語、近代でサ変動詞用法が認められたものの、現代ではその用法が消失・衰退した近代・準近代サ変は 266 語であった。

鄧 (2021) は重要漢語を抽出するため、BCCWJ と「日本語日常会話コーパス」(CEJC) を用いて調査を行った。BCCWJ における 13 の言語変種のうち、韻文を除き、書籍については、「出版・書籍」、「図書館・書籍」及び「特定目的・ベストセラー」を合わせて 1 つの言語変種とみなした。また、CEJC は 1 つの言語変種とみなし、書籍、雑誌、新聞、白書、教科書、広報紙、Yahoo!知恵袋、Yahoo!ブログ、法律、国会会議録、話し言葉 (CEJC) という 11 変種のデータを調査した。具体的な調査手法については、まず、BCCWJ におけるすべての 2 文字の漢語を抽出した。次に、それらの漢語に対して各言語変種における頻度と、各言語変種における語の順位を調べた。最後に、各漢語の重要度を「コーパスにおける出現頻度 (PMW) × (1/言語変種順位平均値)」で計算し、その値に基づき順位付けを行い、重要漢語 2865 語を抽出した。

以下、それぞれの先行研究で抽出された上位 30 語を例として示す。なお、張 (2014) では語の順位付けはなされていないため、サンプルとして、「VN 型-例内-4 級・3 級」に分類される二字漢語動詞 18 語を五十音順で選んで表示している。

表 2 先行研究で抽出された語 (上位 30 語)

先行研究	単語
李(2013) 93 語	存在, 利用, 紹介, 説明, 使用, 確認, 理解, 表示, 発見, 発生, 注目, 指摘, 期待, 実施, 用意, 発表, 注意, 開催, 構成, 評価, 設定, 変化, 提供, 形成, 実現, 表現, 展開, 参加, 検討, 意味
張(2014) 4383 語	(4 級) 結婚, 作文, 授業, 病気 / (3 級) 見物, 講義, 失礼, 出席, 食事, 心配, 水泳, 卒業, 退院, 暖房, 入院, 入学, 発音, 返事
間淵(2016) 通時的 2043 語 準現代・現代 82 語 近代・準近代 266 語	(現代サ変) 機能, 電話, 集約, 絶句, 骨折, 妥当, 同期, 遭難, 建国, 現像 / (準現代サ変) 作業, 補完, 重宝, 一望, 故障, 創業, 異動, 仲介, 視聴, 法定 / (通時的サ変) 利用, 使用, 存在, 説明, 理解, 表示, 確認, 紹介, 選択, 発生
鄧(2021) 2865 語	必要, 関係, 以上, 社会, 情報, 問題, 生活, 地域, 学校, 利用, 会社, 研究, 自分, 環境, 活動, 内容, 企業, 事業, 結果, 基本, 可能, 一般, 方法, 状況, 経済, 調査, 世界, 大学, 現在, 使用

上記を一瞥すると、最上位 30 語と言いながら、選ばれた語はまちまちで、リスト間の重なりがむしろ小さいことがわかる。このことは、使用するデータや手法によって、「重要語」の中身が大きく変わってくることを示唆する。

従来の研究は、様々なデータを駆使して重要な漢語動名詞の特定を試みており、その手法においても、また、最終的に提案されたリストについても、重要な価値を有している。一方で、様々な

言語資源が活用できるようになった現在の時点から見ると、それらには、(1) 話し言葉の分析が少なく、話し言葉コーパスを用いる場合も発話の場面や内容の違いを考慮した調査がなされていない、(2) 書き言葉コーパスについても詳細な変種区分ごとの分析が少ない、(3) 重要語の決定基準は頻度に偏っており、レンジ(全変種のうち、当該語が出現した変種の数・割合)が十分に考慮されていない、(4) 初中級学習者にとって重要な 200 語程度のリストが作成されていない、といった問題も残されている。

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、過去の重要漢語動名詞の抽出研究は大きな成果を挙げたものの、(1) 話し言葉分析、(2) 書き言葉の詳細変種別分析、(3) レンジ情報活用、(4) 初中級学習者リスト作成、といった点で、不足も認められる。そこで、本研究では、(1') 話し言葉コーパスの全面活用、(2') 書き言葉・話し言葉コーパスの詳細変種別分析、(3') レンジと頻度を組み合わせた重要度指標の採択、(4') 初中級学習者向けの 200 語程度の抽出を試みたい。なお、(4') に関しては、各種の漢語語彙表のサイズを参考にしつつ、とくに初中級向けということで 200 語という範囲を暫定的に定めることとした。

重要漢語動名詞の抽出にあたっては、事前に、書き言葉・話し言葉を含む幅広い変種間での漢語動名詞頻度のずれや、各変種で多用される漢語動名詞の内容的一致度を見ていくことが不可欠になるだろう。そこで、本研究では以下の研究設問を設定した。

- RQ1 漢語動名詞全体の頻度は変種間でどの程度異なるか。また、漢語動名詞が多い変種・少ない変種にはどのような共通性があるか？(言語変種別の漢語動名詞全体頻度)
- RQ2 各変種の最頻出漢語動名詞はどの程度の変種間重複度を有するか？(最頻出語の重複度)
- RQ3 高頻度かつ高レンジの漢語動名詞はどのようなものであるか？(重要漢語動名詞の抽出)
- RQ4 上記は日本語の初中級者が学ぶ上で適切な語となっているか？(重要漢語動名詞の教育的妥当性)

3.2 データ

すでに述べたように、本研究では書き言葉・話し言葉を含む現代日本語の多様なジャンル差に配慮して漢語動名詞の使用実態を解明するとともに、汎用的な初中級学習者向けの重要漢語動名詞を抽出していく。そこで、本研究では、「現代日本語書き言葉均衡コーパス(BCCWJ)」と「日本語日常会話コーパス(CEJC)」を主たるデータとして利用する。

BCCWJ は新聞、雑誌、書籍、白書、教科書など 13 種の大ジャンルに区分されており、ジャンルによってはさらに詳細な下位区分がある。本研究では、(1)「韻文」を除外する、(2)「出版・書籍」と「図書館・書籍」、「特定目的・ベストセラー」を書籍にまとめる、(3)「書籍」は日本十進分類法(NDC)による 10 種の内容区分(総記、哲学、歴史、社会科学、自然科学、技術工学、産業、芸術

美術, 言語, 文学)でわけ, (4)「雑誌」は 6 変種, 「白書」は 9 変種, 「教科書」は 10 変種, 「Yahoo!知恵袋」は 13 変種, 「Yahoo!ブログ」は 13 変種の内容区分にわけ, (5)内容的な下位区分のない「新聞」「法律」「国会会議録」「広報紙」などはそのまま 1 変種とみなす, という方針をとる。これにより, 本研究が調査対象とする書き言葉の内容変種は 65 種となる。

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」)によってジャンルに分けられている。本研究は, 「場面」情報を会話内容の分類基準とみなし, CEJC のデータを「自宅」「職場」「学校」「施設」「室内」「屋外」「交通機関」という 7 種にわけ調査する。

以上より, 本研究で調査対象となるのは, 書き言葉と話し言葉を合わせて 72 種の内容変種である。各々の基本データを下記に示す。なお, 個々の変種の総語数については, 『『現代日本語書き言葉均衡コーパス』短単位語数表 (Version 1.1)』と『『日本語日常会話コーパス』モニター公開版の語彙』(2020)を参照する。

表 3 72 の変種の基本データ

変種	総語数	変種	総語数
書籍・総記	1,752,604	知恵袋・インターネット, PC と家電	1,166,139
書籍・哲学	3,826,239	知恵袋・ビジネス, 経済とお金	408,553
書籍・歴史	6,815,084	知恵袋・職業とキャリア	379,506
書籍・社会科学	15,694,683	知恵袋・ニュース, 政治, 国際情勢	439,212
書籍・自然科学	5,226,786	知恵袋・スポーツ, アウトドア, 車	800,557
書籍・技術工学	4,752,183	知恵袋・暮らしと生活ガイド	1,047,234
書籍・産業	2,579,106	知恵袋・健康, 美容とファッション	2,369,297
書籍・芸術美術	4,765,710	知恵袋・子育てと学校	932,555
書籍・言語	1,205,702	知恵袋・マナー, 冠婚葬祭	332,404
書籍・文学	24,082,837	知恵袋・教養と学問, サイエンス	947,806
雑誌・総合	3,859,027	知恵袋・地域, 旅行, お出かけ	460,022
雑誌・教育学芸	524,537	知恵袋・Yahoo! JAPAN	1,407,240
雑誌・政治経済商業	187,737	ブログ・ビジネスと経済	430,174
雑誌・産業	44,655	ブログ・コンピュータとインターネット	105,458
雑誌・工業	639,325	ブログ・生活と文化	1,247,526
雑誌・厚生医療	76,807	ブログ・エンターテインメントと趣味	3,743,547
白書・国土交通	654,715	ブログ・家庭と住まい	671,792
白書・外交	358,150	ブログ・政治	197,689
白書・安全	1,302,641	ブログ・健康と医学	378,401
白書・教育	85,516	ブログ・学校と教育	179,454

白書・環境	337,420	ブログ・科学	112,180
白書・福祉	738,446	ブログ・出会い	35,028
白書・科学技術	509,009	ブログ・地域	592,396
白書・経済	1,233,799	ブログ・芸術と人文	756,723
白書・農林水産	420,652	ブログ・Yahoo!サービス	4,534,707
教科書・国語	162,316	新聞	1,061,729
教科書・数学	104,231	法律	1,206,077
教科書・理科	312,834	国会会議録	5,596,623
教科書・社会	290,266	広報紙	4,620,337
教科書・外国語	28,100	話し言葉・自宅	122,727
教科書・技術家庭	87,565	話し言葉・職場	32,723
教科書・芸術	69,838	話し言葉・学校	53,207
教科書・保健体育	22,274	話し言葉・施設	261,987
教科書・情報	41,333	話し言葉・室内	92,495
教科書・生活	1,357	話し言葉・屋外	20,957
知恵袋・エンターテインメントと趣味	1,086,481	話し言葉・交通機関	25,231

漢語動名詞の頻度情報の取得には、「中納言」を使用する。データの採集は 2021 年 10 月に実施した。

3.3 手法

この節では、研究設問順に研究手法を具体的に述べる。

RQ1(言語変種別の漢語動名詞全体頻度)では、まず、72 種の内容変種ごとに、中納言の短単位検索を用い、「する」を伴い、サ変動詞として機能する用法を持つ 2 文字の漢語動名詞の抽出を行う。2 文字の漢語動名詞に絞った理由としては、2 文字漢語が日本語の漢語語彙の中心的な位置を占めているためである。実際、3 文字以上の漢語のほとんどは 2 文字の漢語が語基になっている(例:再検討, 全面停止)。また、1 文字の漢語は和語の複合要素との結合が高く、ほかの漢語とは文法的振る舞いが大きく異なっている(例:愛する, 信ずる)(間淵, 2016)。中納言の検索条件として、キーは「語種」が「漢」,「品詞」の「小分類」が「名詞-普通名詞-サ変可能」,後方共起は「語彙素読み」が「スル」と設定する。検索したすべてのデータをダウンロードし、2 文字の漢語動名詞でないものを手作業で取り除く。その後、変種ごとに出現するすべての漢語動名詞の総頻度を調査したうえで、100 万語あたりの調整頻度(PMW)に換算して比較を行う。このことにより、漢語動名詞が多い変種・少ない変種を特定し、どのような共通性があるか考察する。

次に、**RQ2**(最頻出語の重複度)では、まず、72 変種の各々において頻度 1 位となる延べ 72 語を抽出し、重複を除く。この作業により、異なり語数は 36 語となった。次に、36 語の各々について、

全 72 種のうち、いくつの変種で 1 位になっているかを調べ、比率値を計算する。

RQ3(重要漢語動名詞の抽出)では、まず、粗頻度が 1 の単語を削除する。これは、「孤語」については偶然の出現である可能性が否定できないためである。次に、該当する漢語の数が 10 語未満となった 4 変種(「話し言葉・屋外(8 語)」、「教科書・生活(2 語)」、「知恵袋・Yahoo! JAPAN(8 語)」、「話し言葉・交通機関(8 語)」)を除く。その後、残った 68 変種からそれぞれ高頻度上位 10 語、延べ 680 語を抽出する。重複を除くと、異なり語数は 214 語となった。続いて、214 語に対して、「平均頻度×レンジ比率(当該語が出現する変種数/全変種数)」という式で重要度を計算し、その値に基づき順位付けを行う。最終的に上位 200 語を初級から中級前半までの学習者が優先的に学ぶべき重要漢語動名詞として抽出する。

重要度の計算の例として、「作成」と「確認」という 2 つの漢語動名詞を取り上げる。「作成」の場合、68 変種における平均頻度は 46.99、出現する変種数は 5 である。ゆえに、重要度は $46.99 \times (5/68) = 3.455$ となる。「作成」の頻度順位は 8 位であったが、この調整により順位は 23 位に下がった。次に、「確認」の場合、68 変種における平均頻度は 27.64、出現する変種数は 19 である。ゆえに、重要度は $27.64 \times (19/68) = 7.722$ となる。「確認」の頻度順位は 26 位であったが、調整により順位は「作成」を上回り、9 位に上昇した。

その後、選定された漢語動名詞について、ジャンル別の傾向を探るため、重要漢語動名詞上位 50 語を第 1 アイテム、68 種のジャンルを第 2 アイテムとする頻度表に対して対応分析を実施する。対応分析とは、頻度表における行・列の関係を組み替え、頻度表に含まれる情報を少数の成分(次元)にまとめることで、行・列を整理する解析法である(石川他, 2010)。

RQ4(重要漢語動名詞の教育的妥当性)では、李(2013)が抽出した 93 語(以下、A リスト)と本研究で得られた 200 語中の最上位 100 語(以下、B リスト)を比較し、B リストが初中級の学習者に適した漢語動名詞を選んでいるか検証する。検証は 2 段階で行う。まず、第 1 段階として、A リストのみ、B リストのみに含まれる語を抽出し、日本語文章難易度判別システム「jReadability」(<http://jreadability.net/>) (李, 2017)を用い、平均難度を調査する。平均難度の計算については、「日本語教育語彙表」の 6 つのレベルと級外レベルを 1~7 の難度値とみなし、両リストの平均値を比較する。

その後、第 2 段階として、A リストのみ、B リストに含まれる語のうち、同一の難度レベルで、かつ、総画数もほぼ同じ漢語動名詞を 10 組選定し、アンケートを作成した(表 4)。提示順については左側が A リスト独自語、右側が B リスト独自語であるが、協力者にはこのことは一切伝えていない。

表 4 アンケートの調査内容

下記の単語のペアのうち、「日本語教育で初級者に先に教えた方がよい、基本的な語である」と思う語を選んでください。				
Q1	決定	<input type="radio"/>	出席	<input type="radio"/>
Q2	評価	<input type="radio"/>	更新	<input type="radio"/>

Q3	強調 ○	登録 ○
Q4	維持 ○	演奏 ○
Q5	解放 ○	我慢 ○
Q6	構成 ○	検索 ○
Q7	区別 ○	反応 ○
Q8	掲載 ○	修飾 ○
Q9	考慮 ○	感染 ○
Q10	確立 ○	促進 ○

これらのペアについて、15名の協力者に「日本語教育で初級者に先に教えた方がよい、基本的な語である」と思うほうを選ばせる。協力者の内訳は、中国語母語の日本語上級学習者10名（いずれも日本語学習歴は5年以上で留学籍がある。学生3名、会社員2名、日本語教師2名、日本語研究者3名）および日本語母語話者5名（いずれも学生）、合計15人である。協力者数は全15名のため、仮にある単語ペアについて、全員がAリストの語を選んだとすると、この語の得点はAリストが15点、Bリストが0点となる。このようにして10種の語の各々について得点を調べ、最終的に全体を平均した。

4. 結果と考察

4.1 RQ1 言語変種別の漢語動名詞全体頻度

書き言葉・話し言葉コーパスにおける72の内容変種別に漢語動名詞全体の出現頻度を調べ、100万語あたりの調整頻度(PMW)に換算したところ、以下の結果を得た。

表5 言語変種別漢語動名詞の使用頻度

#	変種	PMW	#	変種	PMW
1	教科書・情報	30,822.83	37	書籍・言語	10,863.38
2	法律	21,787.17	38	書籍・芸術美術	10,741.53
3	白書・外交	20,812.51	39	国会会議録	10,321.40
4	白書・農林水産	20,282.80	40	ブログ・コンピュータとインターネット	10,155.70
5	白書・環境	19,032.66	41	知恵袋・スポーツ, アウトドア, 車	9,408.45
6	白書・教育	18,873.66	42	知恵袋・教養と学問, サイエンス	9,104.18
7	白書・安全	18,666.69	43	知恵袋・暮らしと生活ガイド	8,804.15
8	教科書・保健体育	18,452.01	44	知恵袋・マナー, 冠婚葬祭	8,628.06
9	教科書・社会	18,279.78	45	雑誌・総合	8,581.18
10	白書・国土交通	17,836.77	46	ブログ・科学	8,201.11
11	教科書・理科	17,760.22	47	雑誌・教育学芸	8,108.10

12	書籍・総記	17,281.14	48	知恵袋・子育てと学校	7,965.21
13	新聞	17,094.76	49	ブログ・学校と教育	7,929.61
14	雑誌・産業	17,041.76	50	ブログ・地域	7,906.87
15	白書・科学技術	16,883.79	51	知恵袋・健康, 美容とフ ァッション	7,792.61
16	書籍・自然科学	16,598.54	52	知恵袋・地域, 旅行, お 出かけ	7,538.77
17	書籍・技術工学	15,913.95	53	教科書・国語	7,405.31
18	雑誌・政治経済商業	15,756.08	54	ブログ・エンターテイン メントと趣味	7,373.22
19	知恵袋・インターネット, PCと家電	15,484.43	55	知恵袋・エンターテイン メントと趣味	7,248.17
20	書籍・産業	15,411.93	56	書籍・文学	7,247.36
21	雑誌・工業	15,169.12	57	ブログ・健康と医学	7,077.15
22	教科書・技術家庭	15,120.20	58	ブログ・Yahoo!サービス	7,072.78
23	白書・福祉	15,065.42	59	ブログ・出会い	6,623.27
24	白書・経済	14,974.89	60	ブログ・芸術と人文	6,600.83
25	ブログ・政治	14,219.30	61	ブログ・生活と文化	6,505.68
26	雑誌・厚生医療	14,178.40	62	ブログ・家庭と住まい	6,205.79
27	書籍・歴史	13,768.58	63	教科書・数学	6,025.08
28	知恵袋・ビジネス, 経済 とお金	13,422.98	64	教科書・生活	5,158.44
29	教科書・外国語	13,096.09	65	話し言葉・自宅	3,006.67
30	書籍・社会科学	13,036.64	66	話し言葉・施設	2,946.71
31	知恵袋・ニュース, 政 治, 国際情勢	12,925.42	67	話し言葉・屋外	2,624.42
32	書籍・哲学	12,305.82	68	話し言葉・職場	2,567.00
33	教科書・芸術	12,213.98	69	話し言葉・室内	2,508.24
34	知恵袋・職業とキャリア	11,659.89	70	話し言葉・交通機関	2,457.29
35	広報紙	11,496.56	71	話し言葉・学校	2,300.38
36	ブログ・ビジネスと経済	10,918.84	72	知恵袋・Yahoo! JAPAN	24.87

表5から分かるように、漢語動名詞の使用頻度は変種別間において一定の異なりがある。以下、詳しく見ていこう。まず、72の変種のうち、漢語動名詞の使用頻度が最も高い変種は「教科書・情報」(PMW:30,822.83)、最も低い変種は「知恵袋・Yahoo! JAPAN」(PMW:24.97)であった。変種間での漢語動名詞の頻度差は最大1,284倍に達する。これにより、漢語動名詞の使用量が内容変種によって大きく異なることが確認された。

次に、72の変種のうち、漢語動名詞が多い変種と少ない変種を整理したい。

表6 漢語動名詞が多い・少ない変種

漢語動名詞が多い変種		漢語動名詞が少ない変種	
変種	PMW	変種	PMW
教科書・情報	30,822.83	教科書・数学	6,025.08
法律	21,787.17	教科書・生活	5,158.44

白書・外交	20,812.51	話し言葉・自宅	3,006.67
白書・農林水産	20,282.80	話し言葉・施設	2,946.71
白書・環境	19,032.66	話し言葉・屋外	2,624.42
白書・教育	18,873.66	話し言葉・職場	2,567.00
白書・安全	18,666.69	話し言葉・室内	2,508.24
教科書・保健体育	18,452.01	話し言葉・交通機関	2,457.29
教科書・社会	18,279.78	話し言葉・学校	2,300.38
白書・国土交通	17,836.77	知恵袋・Yahoo! JAPAN	24.87

表 6 から分かるように、漢語動名詞が最も多い 10 変種は、書き言葉である教科書・法律・白書の 3 ジャンルに集中していた。これらはいずれも、伝達内容に重点があるもので、大量の情報を一方向的かつ効率的に伝達するジャンルでは漢語動名詞の出現頻度が上昇すると言える。一方、漢語動名詞が最も少ない 10 変種には、話し言葉の全変種と、低学年向けの教科書、知恵袋などのジャンルが含まれる。これらはいずれも、読み手を意識して内容をわかりやすく伝えるもので、情報を双方向的に伝達し、読み手・聞き手との共感醸成を重視するジャンルでは漢語動名詞が出現しにくいと言える。

4.2 RQ2 変種別最頻出語の重複度

各変種の最上位語（36 種）について、変種間での重複度（72 種中、いくつの変種で最上位語になっているか）を調査したところ、以下の結果を得た。

表 7 変種別最頻出語の重複度の内訳

重複度	語数・漢語 VN
11.1% (8/72 変種)	1 語 (実施)
9.7% (7/72 変種)	1 語 (利用)
8.3% (6/72 変種)	2 語 (紹介, 参加)
5.6% (4/72 変種)	1 語 (存在)
4.2% (3/72 変種)	1 語 (表示)
2.8% (2/72 変種)	8 語 (勉強, 表現, 発表, 入院, 説明, 使用, 購入, 結婚)
1.4% (1/72 変種)	22 語 (連絡, 輸出, 放送, 妊娠, 投票, 導入, 定着, 探検, 卒業, 増加, 整理, 出品, 撮影, 検討, 怪我, 計算, 規定, 感染, 加入, 改善, 開催, 演奏)

表 7 より、重複度は最高でも 1 割程度にとどまることが明らかになった。また、72 変種中、1 変種でしか最上位語になっていない語は 22 種で、異なり語数 36 種の 6 割以上を占めることもわかった。これにより、漢語動名詞は、最上位語に限っても、想像以上に変種間

の開きが大きいことが改めて確認された。このことは、重要漢語動名詞抽出において、言語データをかたまりとして処理するのではなく、できる限り細かい変種ごとに頻度調査を行い、重要語を抽出しようとする本研究の方向性が妥当であることを傍証するだろう。

4.3 RQ3 重要漢語動名詞の抽出

漢語動名詞の出現が少なかった 4 変種を除く 68 変種について、それぞれ高頻度上位 10 語、合計 680 語を選び、重複を除いた異なり 214 語に対して、68 変種における平均頻度とレンジ比率を掛け合わせる「重要度」を計算したところ、初級者向けの重要漢語動名詞 200 語を抽出した。リストの全体は本論文の末尾に掲載する。以下は、上位 5 語のコーパス用例である。なお、a は書き言葉、b は話し言葉用例である。当該の漢語動名詞は【】記号で示す。

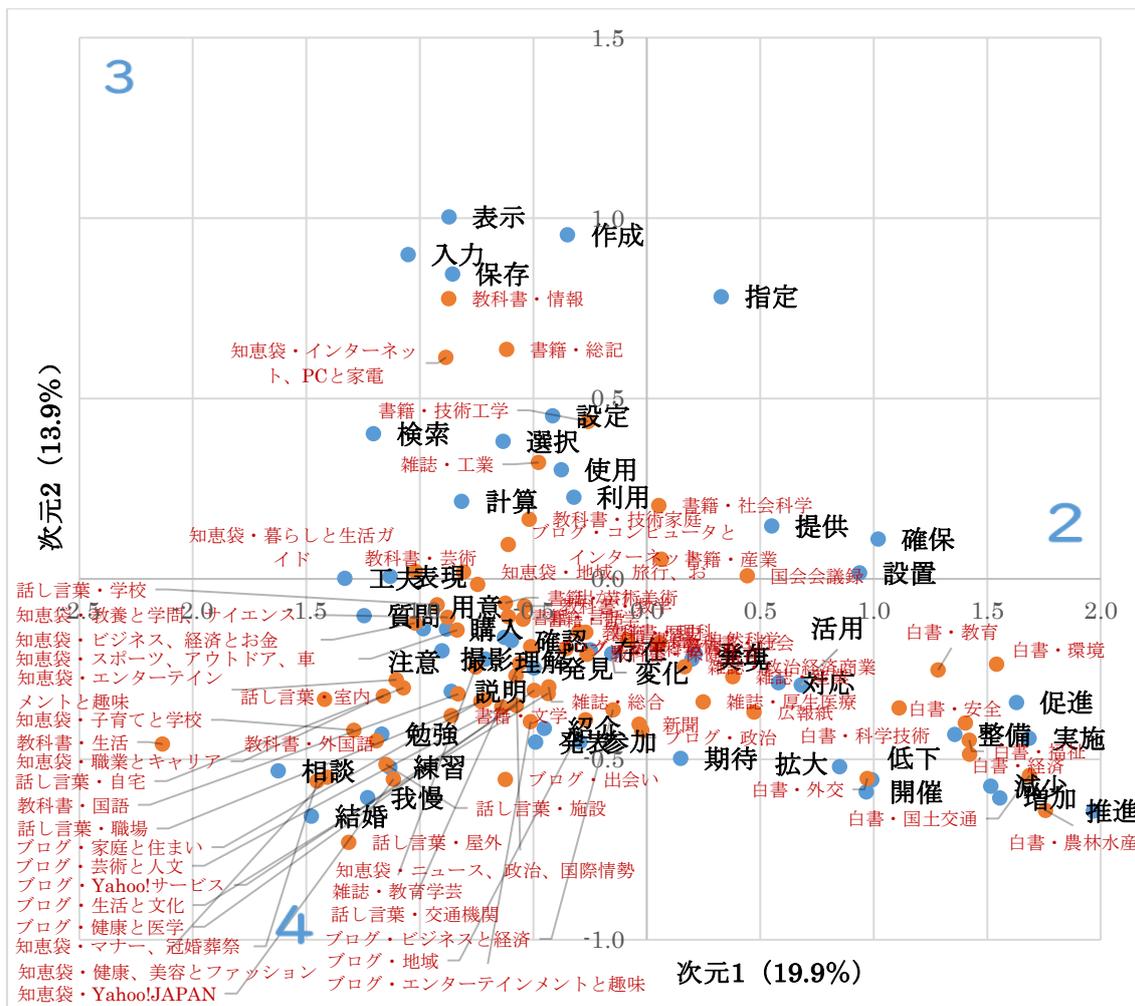
- (9a) ブラウザー翻訳の上下対訳を【利用】すると、表示レイアウトを崩さずに原文と日本語を同時に表示できる。(雑誌・工業, PM35_00179)
- (9b) いや、だからそういう生体のそういう機能を色々なあの工学的に【利用】してるわけですよ(話し言葉・学校, T006_001)
- (10a) さらに、農協におけるホームヘルパー及び介護活動コーディネーターの養成を引き続き【実施】したほか、農協高齢者介護活動の本格的実施のためのマニュアルの策定、モデルプランの作成に対して新たに支援した。(白書・農林水産, OW5X_00090)
- (10b) こっちは雪の時にえっとわたしが勝手に判断して強行【実施】をしたら遅刻して、全員ね(話し言葉・施設, T009_021b)
- (11a) フェントン・ブレスラーの『遺言状のたのしみ』(石神力訳)という本には幽明界を異にした人間の悲喜劇が興味深く【紹介】されている。(書籍・社会科学, LBe3_00023)
- (11b) きょう翔子さんて【紹介】したらさえー最初すぐ思い出せなかったけど(話し言葉・室内, S001_015)
- (12a) 誤解の無いように申し上げますが、ウチはあくまでも、このようなサンプル動画を、ネット技術の研究目的で【使用】しています。(ブログ・家庭と住まい, OY05_02259)
- (12b) でプラス大会参加者もコートも【使用】してもらってあーのまあ会員さんと一緒にテニスやるってゆう雰囲気味わってもらいたい(話し言葉・職場, S001_018)
- (13a) コンピューターを活用した学習システムの構築を図ったり、大学間の連携をデジタル化で【推進】しようという『未来の大学』像を探る組織で、千九百九十九(平成十一)年四月に発足した。(書籍・技術工学, LBp5_00046)

上位 5 語のうち、話し言葉で使用されていない漢語動名詞は「推進」のみであった。漢語

動名詞は一般的に書き言葉的で、話し言葉での使用は制約されると考えられがちだが、上記の用例は、漢語動名詞が話し言葉でもごく普通に使用されることを示している。とくに (2b) の「実施」などは、直観では話し言葉での使用に向かないと思われる語だが、実際には自然な形で使用されている。このことから、今回の分析では、日本語の書き言葉・話し言葉の両面において一般的に使用される漢語動名詞が選定されていると考えられる。

なお、本研究で得られた重要漢語動名詞については、重要度の順に学習していくことが一般的に言うところと望ましい。ただし、学習者が特定の分野の日本語の習得に興味を持っている場合は、重要漢語動名詞のうち、当該分野に特徴的な語を優先的に学ぶという方向も考えられる。こうした目的をふまえ、上位 50 語を対象に対応分析を実施したところ、以下の結果を得た。なお、第 1 象限の上方に、全体から大きく離れて法律ジャンルと漢語動名詞「規定」のみが布置されていたため（これを第 1 グループと称する）、以下ではそれらを除いた部分を示す。また、下図では、筆者の判断でまとまりを判断し、第 2～第 4 グループの指定を行っている。

図 1 対応分析の結果



まず、第1次元によって全体は左右に2分割された。右側には、白書・国会会議録・社会科学や産業に関わる書籍など、フォーマルで、かつ、社会科学的（経済・政治など）なジャンルが布置されている。一方、左側上部にはインターネット、PC（コンピュータ）、技術工学、情報工学など、コンピュータ工学系ジャンルがまとまっている。また、左側下部には、ブログや知恵袋、話し言葉といったインフォーマルで雑多なジャンルが集まっている。ここで扱われる内容は、芸術・趣味・ファッション・学校・ビジネス・スポーツ・出会いなど、幅広い。

一般に、漢語動名詞は書き言葉・フォーマル的なジャンルで多用されると思われるが、インフォーマルなジャンルと強く結びつく漢語動名詞もあることが確認された。下記はジャンル別の主要漢語動名詞のリストである。

表8 ジャンル別重要漢語動名詞

ジャンル	[1] 法律	[2] 社会科学	[3] コンピュータ工学	[4] 一般
特徴語	規定	確保, 設置, 推進, 実施, 整備ほか	入力, 保存, 作成, 表示, 検索ほか	購入, 変化, 紹介, 練習, 我慢ほか

重要度順リストに加え、こうした分野別リストを用意しておくことで、学習者はそれぞれの学習ニーズに応じて、より体系的に漢語動名詞の学びを進めることができるであろう。

4.4 RQ4 重要漢語動名詞の教育的妥当性

最後に、本研究で抽出された重要漢語動名詞の内容的妥当性と独自性を検証するため、李（2013）の重要93語（Aリスト）と、本研究で得られた重要漢語動名詞200語中の上位100語（Bリスト）を比較したところ、以下の結果となった。

表9 李（2013）との比較

Aリストのみ 35語	Bリストのみ 42語	A/B共通 58語
構成, 評価, 要求, 判断, 決定, 認識, 予想, 報告, 発売, 強調, 発展, 開発, 否定, 反映, 代表, 掲載, 支配, 実行, 解決, 解放, 集中, 発揮, 一致, 無視, 維持, 適用, 確立, 分類, 安定, 心配, 重視, 考慮, 販売, 限定, 区別	推進, 質問, 検索, 工夫, 入力, 相談, 整備, 低下, 促進, 練習, 確保, 計算, 撮影, 我慢, 推移, 登録, 応援, 改善, 妊娠, 演奏, 準用, 上昇, 出席, 帰宅, 活躍, 電話, 該当, 派遣, 助成, 更新, 違反, 反応, 運転, 感染, 離婚, 修飾, 批判, 削除, 供与, 輸出, 相当, 加入	存在, 利用, 紹介, 説明, 使用, 確認, 理解, 表示, 発見, 発生, 注目, 指摘, 期待, 実施, 用意, 発表, 注意, 開催, 設定, 変化, 提供, 形成, 実現, 表現, 展開, 参加, 検討, 意味, 対応, 主張, 完成, 成功, 設置, 作成, 選択, 開始, 採用, 登場, 成立, 導入, 意識, 移動, 指定, 増加, 提出,

		規定, 結婚, 変更, 拡大, 比較, 購入, 記載, 勉強, 観察, 保存, 想像, 活用, 減少
--	--	--

B リストのみの 42 語には, 初級の学習者が学校場面や学習場面で遭遇しやすい「質問」「相談」「練習」「出席」「帰宅」といった語のほか, 「撮影」「応援」「電話」「運転」などの日常生活に関わる語, また, 「工夫」「我慢」「批判」「削除」などの基本的な行為にかかわる語が含まれていることがわかった。これらの語が選ばれたことは, 本リストが, 一定の範囲で, 初級者にとって必要な語彙をうまく選定できていることを示唆するが, この判断は主観的なもので, 何らかの外的基準による検証が必要である。

そこで, 以下, B リストの教育的妥当性を 2 段階で検証したい。まず, 第 1 段階として, A リストのみの 35 語と, B リストのみの 42 語の平均難度と, 難度レベル別の語数を比較したところ, 以下の結果を得た。

表 10 李 (2013) と本研究リストの難度比較

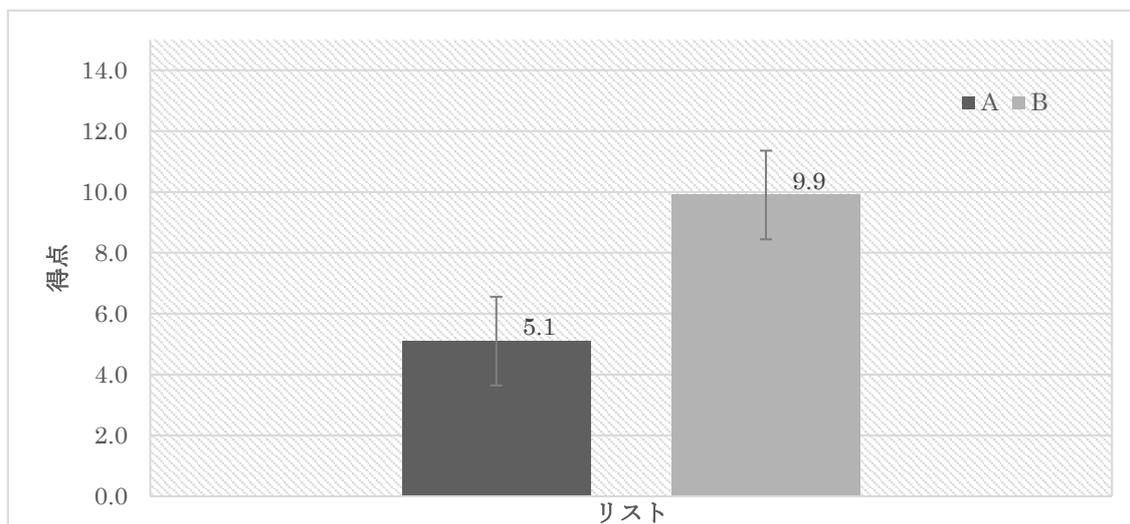
	A リストのみ 35 語	B リストのみ 42 語
難度スコア	4.00	4.07
初級前半 (1)		1 語 (電話)
初級後半 (2)	1 語 (心配)	2 語 (質問, 練習)
中級前半 (3)	3 語 (決定, 発売, 代表)	9 語 (相談, 低下, 計算, 応援, 出席, 帰宅, 違反, 運転, 輸出)
中級後半 (4)	26 語 (構成, 評価, 要求, 判断, 予想, 報告, 強調, 発展, 開発, 否定, 反映, 支配, 実行, 解決, 解放, 集中, 一致, 無視, 維持, 適用, 分類, 安定, 重視, 販売, 限定, 区別)	17 語 (工夫, 入力, 検索, 撮影, 整備, 我慢, 登録, 改善, 活躍, 更新, 演奏, 上昇, 反応, 離婚, 削除, 相当, 加入)
上級前半 (5)	5 語 (認識, 掲載, 發揮, 確立, 考慮)	10 語 (推進, 促進, 推移, 確保, 妊娠, 該当, 派遣, 感染, 修飾, 批判)
級外 (7)		3 語 (準用, 助成, 供与)

難度スコアに着目すると, B リストのほうがやや高いという結果になったが, 一元配置の分散分析を行ったところ, 差は有意ではなかった ($F(1,75)=.098$, $\eta_p^2 = .001$, $p=.755$, ns)。つまり, 難度の点で, 両リストは同等となり, B リストのほうが平易な語を選んでいるわけではないことが示された。

そこで, 第 2 段階の検証として, 日本語学習者・教師・日本語母語話者 15 名に対するアンケート調査を行い, A, B リストから抽出された単語ペアのうち, 「日本語教育で初級者

に先に教えた方がよい、「基本的な語」を問うたところ、結果は以下の通りとなった。下図の得点は、10種の調査語に対して、15人の協力者のうち何人がその語をより基本的であると答えたかを示す。

図2 「基本的な語」と回答した人数の平均値（全15人中）



注：エラーバーは標準誤差（2.06）を示す。

図2より、Bリストのほうが、基本的であると判断された語が多い可能性が示された。 t 検定を実施した結果、先行リストと本研究リストの間の差は有意であることが確認された（ $t(18)=2.060$, $p=.032$, $d=.998$, 効果量は「大」）。これにより、本研究が抽出した重要漢語動名詞リストは、平均難度の点では、先行リストとはっきりした差がないものの、日本語学習者や母語話者の直観においてより基本的と感じられる語を多く含むものとなっていることが確認された。

5. まとめ

以上、本研究では、BCCWJとCEJCに含まれる言語データを72の変種に分類し、変種ごとに漢語動名詞の出現状況を網羅的に調査し、多様な現代日本語における漢語動名詞の使用実態を解明するとともに、初中級学習者向けの汎用的な重要漢語動名詞200語を抽出した。

RQ1（変種別頻度）では、現代日本語書き言葉・話し言葉の各変種間における漢語動名詞の頻度の差は大きく、1,000倍以上になる場合もあること、また、漢語動名詞は大量の情報を効率的に伝達することを優先する変種で多く、話し言葉や、低学年向けの教科書などのような双方向的に情報を伝達し、読み手・聞き手間の共感醸成を重視する変種で少なくなることが確認された。

RQ2（最頻出語の重複度）では、変種別の最上位漢語動名詞の重なり度合いは想像以上に

低いことが確認された。

RQ3 (重要漢語動名詞の抽出) では、上記をふまえ、「平均頻度×レンジ比率」という尺度に基づき、重要漢語動名詞 200 語を抽出した。なお、上位 50 語を対象に対応分析を実施して、分野別の主要漢語動名詞のリストを作成した。

RQ4 (重要漢語動名詞の教育的妥当性) では、本研究で抽出された重要漢語動名詞の内容の妥当性と独自性を 2 段階で検証したところ、本リストには、学習者・母語話者がともに「基本的」であると感じる初級者向けの重要漢語動名詞が盛り込まれていることが確認された。

今後は、抽出された初中級学習者向けの重要漢語動名詞 200 語に基づき、漢語動名詞の語彙的・統語的振る舞いを考察するとともに、研究で得られた成果を具体的な教材や教授法の開発に生かす方向を考察していきたい。

引用文献

- 浅原正幸・松本裕治(2003)『ipadic version 2.7.0 ユーザーズマニュアル』奈良先端科学技術大学院大学情報科学研究科自然言語処理学講座。
- 石川慎一郎・前田忠彦・山崎誠(2010)『言語研究のための統計入門』くろしお出版。
- 大村舞・柏野和佳子・山崎誠(2020)『『日本語日常会話コーパス』モニター公開版の語彙』国立国語研究所。
- 影山太郎(1993)『文法と語形成』ひつじ書房。
- 国立国語研究所コーパス開発センター(2021)『『現代日本語書き言葉均衡コーパス』短単位語数表 (Version 1.1)』国立国語研究所。
- 張厚泉・許小明(2009)『新日本語教程』人民教育出版社。
- 張志剛(2014)『現代日本語の二字漢語動詞の自他』くろしお出版。
- 伝康晴・山田篤・小椋秀樹・小磯花絵・小木曾智信(2008)『UniDic version 1.3.9 ユーザーズマニュアル』(https://ccd.ninjal.ac.jp/unidic/UNIDIC_manual.pdf)
- 鄧琪(2021)「非和語系語彙の連体修飾用法に対する一考察:異なる連体修飾標識の許容度に注目した計量調査」『統計数理研究所共同研究レポート』444,123-138。
- 平尾得子(1990)「サ変動詞をめぐる」『待兼山論叢.日本語学篇』24,57-73。
- 間淵洋子(2016)「近現代漢語におけるサ変動詞用法の変化:形態論情報付きコーパスを用いて」『国際日本語学研究論集』4,17-36。
- 山田孝雄(1940)『國語の中に於ける漢語の研究』明治書院。
- 李在鎬(編著)(2017)『文章を科学する』ひつじ書房。
- 李楓(2013)「現代日本語における汎用的漢語サ変動詞の抽出とその内部構成の検討」『第 4 回コーパス日本語学ワークショップ予稿集』.101-110。

付表 現代日本語重要漢語動名詞リスト(全 200 語)

順位	漢語 VN	重要度	順位	漢語 VN	重要度	順位	漢語 VN	重要度
1	利用	74.396	68	観察	0.279	135	発展	0.053
2	実施	40.524	69	開始	0.272	136	制作	0.053
3	紹介	27.240	70	比較	0.266	137	出産	0.052
4	使用	20.262	71	形成	0.233	138	学習	0.052
5	推進	12.057	72	変更	0.222	139	就職	0.050
6	増加	11.889	73	出席	0.218	140	募集	0.049
7	参加	11.653	74	帰宅	0.200	141	定植	0.048
8	購入	11.595	75	活躍	0.195	142	出演	0.048
9	確認	7.722	76	電話	0.190	143	整理	0.048
10	説明	7.615	77	該当	0.176	144	反映	0.047
11	存在	6.776	78	完成	0.169	145	転職	0.046
12	発表	6.769	79	派遣	0.161	146	告白	0.044
13	質問	6.155	80	助成	0.156	147	合格	0.043
14	表示	5.846	81	想像	0.154	148	下落	0.043
15	対応	5.125	82	更新	0.150	149	発売	0.041
16	結婚	5.035	83	違反	0.146	150	発行	0.041
17	表現	4.712	84	反応	0.143	151	由来	0.040
18	発生	4.475	85	成立	0.140	152	抽出	0.039
19	減少	4.453	86	意識	0.134	153	変形	0.039
20	開催	4.240	87	運転	0.126	154	冷凍	0.039
21	理解	3.782	88	感染	0.117	155	記入	0.038
22	注意	3.493	89	離婚	0.114	156	投票	0.035
23	作成	3.455	90	修飾	0.108	157	診断	0.034
24	期待	2.955	91	導入	0.105	158	投与	0.034
25	設定	2.504	92	意味	0.104	159	分泌	0.034
26	検索	2.489	93	批判	0.103	160	支配	0.034
27	設置	2.420	94	削除	0.101	161	建設	0.033
28	工夫	2.365	95	供与	0.100	162	卒業	0.033
29	入力	2.360	96	記載	0.098	163	希望	0.032
30	選択	1.988	97	輸出	0.097	164	出発	0.031
31	用意	1.813	98	相当	0.097	165	悪化	0.031
32	相談	1.637	99	加入	0.096	166	吸収	0.031
33	勉強	1.534	100	成功	0.096	167	逮捕	0.031

34	活用	1.359	101	適用	0.093	168	怪我	0.030
35	保存	1.257	102	栽培	0.092	169	報道	0.029
36	提供	1.196	103	起動	0.089	170	報告	0.028
37	変化	1.138	104	上達	0.087	171	旅行	0.028
38	規定	1.051	105	發揮	0.087	172	当選	0.028
39	整備	0.978	106	入院	0.084	173	反発	0.027
40	発見	0.941	107	採択	0.083	174	優勝	0.027
41	指定	0.908	108	予約	0.078	175	強調	0.026
42	低下	0.877	109	到着	0.078	176	定着	0.026
43	促進	0.745	110	開発	0.077	177	受診	0.024
44	練習	0.745	111	交換	0.077	178	解決	0.023
45	確保	0.699	112	経過	0.076	179	挨拶	0.022
46	計算	0.688	113	表明	0.075	180	支持	0.022
47	撮影	0.671	114	注文	0.074	181	発音	0.021
48	我慢	0.533	115	招待	0.074	182	活動	0.020
49	実現	0.493	116	準備	0.072	183	化粧	0.020
50	拡大	0.485	117	接続	0.070	184	発注	0.020
51	推移	0.480	118	退職	0.070	185	混雑	0.019
52	展開	0.469	119	服用	0.068	186	宿泊	0.019
53	採用	0.443	120	販売	0.068	187	後悔	0.019
54	登録	0.440	121	強化	0.068	188	参拝	0.019
55	登場	0.435	122	出荷	0.063	189	収録	0.019
56	提出	0.424	123	証明	0.062	190	改正	0.018
57	指摘	0.406	124	作曲	0.062	191	調整	0.018
58	応援	0.406	125	測定	0.062	192	出土	0.017
59	移動	0.405	126	放送	0.061	193	判断	0.017
60	注目	0.353	127	発達	0.059	194	議論	0.017
61	改善	0.340	128	合意	0.059	195	集中	0.016
62	妊娠	0.338	129	達成	0.058	196	終了	0.014
63	検討	0.305	130	排出	0.058	197	否定	0.014
64	演奏	0.300	131	搭載	0.056	198	貢献	0.014
65	主張	0.299	132	申告	0.056	199	成長	0.014
66	準用	0.298	133	沸騰	0.054	200	退治	0.013
67	上昇	0.284	134	実行	0.053			