

BCCWJ と CEJC を用いた「は」と「が」の出現状況調査 —外国人学習者のための日本語指導の指針を求めて—

李 宏宇（神戸大学大学院国際文化学研究科）

An Investigation into the Occurrences of “Wa” and “Ga” Using BCCWJ and CEJC -Seeking Pedagogical Guidelines for Learners of Japanese as a Foreign Language-

Hongyu LI (Graduate School of Intercultural Studies, Kobe University)

要旨

本研究の目的は、外国人学習者のための「は」と「が」の指導方針を考える準備作業として、日本語の様々な言語変種における「は」と「が」の使用実態を記述的に調査することである。BCCWJ と CEJC を用い、計 7 変種を対象として、「は」と「が」の出現頻度と主要機能を分析した。その結果、7 変種中 6 変種において「は」の使用頻度が「が」を上回り、書き言葉では「は」、話し言葉では「が」と親和性が相対的に高いことが確認された。また、書き言葉では「は」の「取り立て」・「対比」の機能、話し言葉では「は」の「主題」の機能と「が」の「主体」・「対象」の機能が相対的に多く用いられることが解明された。以上の結果から、「は」と「が」の使い分けは、文法的要因のみならず、言語変種、談話機能とも密接に関連していることが示された。本研究の知見は、日本語学への理論的貢献にとどまらず、日本語教育における「は」と「が」の指導への応用も期待される。

1. はじめに

国際交流基金(2026)によれば、日本語学習者数ははじめて 400 万人を超え、効果的な日本語コミュニケーション指導への期待が高まっている。膠着語である日本語の理解・産出において鍵を握るものは助詞である。助詞は文中における語と語の文法的関係や談話上の機能を示す重要な役割を持つが、日本語の助詞は種類が多く、用法も複雑であるため、学習者にとって習得困難な項目の一つとされている。

多様な日本語助詞の中で、特に重要と思われるのは、取り立ての「は」と主格表示の「が」である。韓国語を除く主要言語において「は」と「が」に相当する助詞は存在しないため(庵, 2016)、これらの習得は特に困難となりうる。以下、「多言語母語の日本語学習者横断コーパス」(International Corpus of Japanese as a Second Language, I-JAS)に収録された中国語母語(CCM, CCH)およびドイツ語母語(GDE)の学習者によるストーリーライティング作文(SW: 複数コマの漫画を見てストーリーを作文する)の実例を見てみよう。なお、学習者コードの末尾の記号は、日本語能力試験(JLPT)の級を示す。

- (1) ケンは梯子 が→を 使う...(I-JAS, CCM53-SW2-00040-K, N4)
- (2) 犬がバスケットから出て、ケンとマリ が→は びっくりしました。(I-JAS, CCH17-SW1-00020-K, N2)
- (3) でも、二人は地図を見ている時に、犬 は→が バスケットに入りました。(I-JAS,

CCH17-SW1-00020-K, N2)

- (4) 二人とも は→が 料理の準備に夢中になっていて...(I-JAS, GDE04-SW1-00030-K, N1)

注目すべきは、「は」と「が」の誤用が N4 など、日本語を学び始めた初級者のみならず、N2 や N1 といった中上級者にも見られることである。

では、このような誤用が長く残ることの背景として、学習者はそれぞれの母国においてどのような指導を受けているのであろうか。ここでは、日本語学習者数が最多の中国に注目する。表 1 は、中国で広く使用されている『综合日语 第一册 第三版』『新经典日本語 基础教程 第一册』における「は」と「が」の記述の一部である。

表 1 中国の教科書の記述

	「は」	「が」
综合日语 (総合日本語)	<N1 は N2 です>助词「は」提示主语或话题, 读作“wa”。日本語訳: 助詞「は」は主語または話題を提示し、「wa」と発音する。(综合日语,第 2 課ユニット 1, p.23) <は(対比)>表示对比。将两个事物进行对比时, 要比较的事物用助词「は」提示。多使用「～は～が、～は～」的句式。日本語訳: 対比を表す。二つの物を対比させる場合、比較する対象は助詞「は」によって提示される。「～は～が、～は～」という文型が多く用いられる。(综合日语,第 4 課ユニット 2, p.83)	<が(主体: 主語、疑問)>表示句子的主語。接续为名词+が。疑问词作主語时, 要用「が」表达未知的信息, 回答也必须用「が」。日本語訳: 主語を表す。接続は「名詞+が」である。疑問詞が主語となる場合、未知の情報を表すため「が」を用いなければならない。その応答においても必ず「が」を用いなければならない。(综合日语,第 4 課ユニット 1, p.74) <が(主体)>表示客观事件(如现象描写)的主語。接续为名词+が。日本語訳: 客観的な出来事(現象の描写など)を表す主語。接続は「名詞+が」である。(综合日语,第 5 課ユニット 1, p.104)
新经典日本語 (新經典日本語)	<N1 は N2 です>「は」(读作“wa”)是提示助词, 提示主题。日本語訳: 「は」(「wa」と発音する)は取り立て助詞であり、主題を提示する。(新经典日本語,第 1 課, p.11)	
	<N1 は N2 が A いです/AN です>该句型用语描述人或事物的某方面特征。「は」提示句子的主题, 「が」表示句子的主体。日本語訳: 本文型は人や事物の特定側面における特徴を描写する際に用いられる。「は」は文の主題を提起し、「が」は主語を示す。(新经典日本語,第 5 課, p.70)	

最低限の情報は提示されているものの、これだけでは、幅広い状況や文脈において、学習

者が自信を持って「は」と「が」を使い分けることは難しいように思える。このような現状を考えた場合、書き言葉と話し言葉を含む多様な言語変種において、「は」と「が」が実際にどの程度使用され、どのような意味を有しているのかを実証的に調査することが不可欠であろう。そこで本研究は、外国人学習者のための「は」と「が」の指導方針を考えるための準備として、日本語の様々な言語変種における「は」と「が」の使用実態の解明を目指す。

2. 先行研究

日本語における係助詞「は」と格助詞「が」の意味・用法や使い分けについては様々な研究が行われてきたが、本節では、日本語学、コーパス言語学、および第二言語習得研究に基づく研究を中心に概観する。また、日本語研究の方法論として、言語変種を区分した分析の重要性に関する研究についても整理する。

2.1 日本語学における「は」と「が」

まず、日本語学の研究において、野田(1996)は、「は」と「が」の基本的な性質について、係助詞「は」が主に「主題」や「対比」という格を表さない機能を、格助詞「が」が主に「主格」や「排他」という格を表す機能を持っていると指摘している。

野田(1991)は、格助詞は動詞と名詞の格関係を表すものであり、具体的には名詞が動詞の表す動作の主体や対象を示す役割を担っていると指摘した。また、庵(2016)によると、「主題」は文頭に位置して「～について言えば」という意味を表し、その文で述べる内容を提示する成分であるとされ、一方の「主語」はガ格でマークされ、直接目的語ではない名詞句として規定されている。野田(1996)は、「主題をもてるかどうか」「主題をもつかどうか」「何の主題にするか」「主題を明示するかどうか」「どうとりたてるか」という5つの原理を提案し、「は」と「が」の使い分けを明らかにした。「は」と「が」の機能については、「は」が表す主題・対比の機能と、「が」が表す主格・排他の機能という二面性が重視される。さらに、「は」と「が」の対立についても、主題提示と取り立てという二面性から捉えられている。しかし、高見ほか(2004)が指摘するように、「は」と「が」の使い分けは、初級・中級・上級のいずれの日本語学習者にとっても困難である。多くの学習者は、「は」文と「が」文の違いを「単に強調するか否か」という基準のみで区分している。また、「は」が主題を表すのか対比を表すのかは明確に区分できないことが多いため、学習者がこれらを混同し、発信時の誤用につながるケースが散見される。庵(2018)では、「は」と「が」が用いられる環境が以下のように整理されている。「は」は、(1)格成分の後、(2)接続助詞の後、(3)述語の中、などの環境で使用される。一方、「が」は、(1)動作・出来事・状態・存在の主体、または感情の持ち主、(2)感情や生理現象の対象、(3)項の増減を伴わないボイス表現、(4)「ハーガ構文」における「が」名詞句、などの環境で用いられる。日本語記述文法研究会(2009)では、文および談話における「は」と「が」の使い分けが詳細に説明されている。まず「文」においては、(1)文が主題を持ち、その主語が主題として明示される場合は「は」で示され、主題を持たない場合は「が」で示されること、(2)主題が明示される場合は「は」が用いられ、明示されない場合は「は」が用いられないこと、(3)主語を対比的に取り立てる場合は「は」で示され、排他的に取り立てる場合は「が」で示されること、が明らかにされている。「談話」においては、(1)話し手が、聞き手にとって指示対象が特定可能であると判断する場合は主題として「は」で示し、特定不可能であると判断する場合は主題としては示さず「が」で示すこと、(2)人や事物の属性を述べる文では、その人や事物を主題として「は」で提示

し、新たな出来事の生起を述べる文には主題が表れず、主語が「が」で示されること、(3)他の文に従属する文には主題を持たないものがあること、(4)主語を対比的に取り立てる場合は「は」で示され、排他的に取り立てる場合は「が」で示されること、が解明されている。

2.2 コーパス言語学における「は」と「が」

次に、コーパス言語学の研究について、石川(2013)では、「現代日本語書き言葉均衡コーパス」(Balanced Corpus of Contemporary Written Japanese, BCCWJ)より無作為抽出した 1000 例を対象とし、「は」と「が」の選択傾向を計量的視点から質的に検証した。その結果、日本語の書き言葉全般において「は」が優先されるわけではないこと、およびテキストのレジスターなどの非言語的属性が「は」と「が」の選択に一定の影響を及ぼすことが示唆された。杉村(2010)では、魯東大学会話コーパスを用いて、中国語を母語とする日本語学習者の会話データにおける格助詞の誤用の原因が解明されている。その結果、「が」の誤用が多く、特に「を」や「に」を使うべき時に「が」を使い、「が」を使うべき時に「を」を使用する誤用が多く見られる。母語である中国語の影響や、学習者において「は」と「が」の意味や使い分けの理解が不十分であることが原因として考えられる。庵(2019)は、「日本語話し言葉コーパス」(Corpus of Spontaneous Japanese, CSJ)の CSJ-RDB Version1.0 を用いて、三上章による「は」と「が」に関する論説を量的に検証し、日本語教育における重要性を論じている。出現頻度の観点からは、「は」が「文末」や「並列節」において顕著に現れる一方、「が」は「ト節・カ節」「連用節」「連体節」で優位に観察される。また、文節間距離の観点では、「は」は「が」と比較して全体的に長距離の係り受け関係を持ちやすく、その傾向は特に文末において顕著であることを明らかにしている。庵(2021)は、CSJ-RDB Version2.0 を用いて、「は」と「が」の使い分けのフローチャートの妥当性を計量的に検証している。また、その結果を踏まえ、日本語教育の観点からは、教師自身が「は」における「主語を兼ねる場合」と「主語ではない場合」の存在、および「が」が原則として主語を表すという性質を理解し、これらを学習者へ適切に提示していくことが求められる。

2.3 第二言語習得研究における「は」と「が」

さらに、第二言語習得研究について、ゲオルギエバ(2012)は、日本語母語話者が文脈の解釈に基づき「は」と「が」を柔軟に使い分け、同一箇所にも両助詞を使用可能とみなすのに対し、日本語学習者は隣接する言語環境に依存して使い分ける傾向が強く、双方の助詞の許容度が低いと指摘した。その結果、学習者は母語話者と異なり、「は」と「が」の使い分けを極めて二者択一的なものとして捉えていることが明らかにされている。小口(2017)は、上級日本語学習者の談話における「は」と「が」の運用実態を調査し、学習者が多様なストラテジーを用いながら部分的に母語話者に近い使い分けを示しつつも、未出・既出といった言語知識に基づく自律的な選択の段階には至っていないことを明らかにした。この知見に基づき、日本語教育における両助詞の指導において、一定の長さを持つ談話教材を活用して文脈を重視した活動を行うとともに、運用の自動化を促すアプローチが効果的であると提言している。さらに、中西(2024)は、「は」と「が」の習得実態として、学習者は単文レベルでは比較的正しい選択が可能であるものの、複文を含む複雑かつ長大な文章構造においては誤用が増加する傾向にあることを示した。また、この誤用傾向は学習者の習熟度によって異なり、短文レベルではその差が顕著に現れない一方で、複雑な長文においては習熟度の低い学習者の誤用率が高まるという特徴を持つと述べている。朱(2024)では、中国の大学で日本

語を専攻する3、4年生を対象に、アンケート調査を通じて「は」と「が」の習得実態を考察した。その結果、「は」と「が」の使い分けにおける誤用の原因として、主題と主語の概念があいまいであること、機能や用法の理解が不明瞭であること、「は」や「が」の省略・脱落誤用傾向が顕著であることを指摘している。

2.4 日本語研究におけるモード・ジャンル・場面の重要性

最後に、日本語におけるモード・ジャンル・場面に関する研究について、奥野・ディアンニ(2015)、滝島(2020)においては、書き言葉と話し言葉における語彙使用の差異があると指摘されている。近年、BCCWJと「日本語日常会話コーパス」(Corpus of Everyday Japanese Conversation, CEJC)を併用した研究が増加しており、さらにI-JASや「北京日本語学習者縦断コーパス」(Beijing Corpus of Japanese as a Second Language, B-JAS)といった学習者コーパスを用いた研究も益々盛んになってきている(李, 2020)。丸山(2015)は、BCCWJおよびCSJを用いて、書き言葉と話し言葉の違い、およびレジスターごとの違いを分析した。その結果、調査対象となった助詞の中で格助詞の割合が最も高いことが示された。また、BCCWJにおいてはレジスターごとに助詞の使用頻度が異なること、CSJにおいては助詞の使用に話者特性よりも場面による影響が顕著であること、さらにBCCWJの「白書」・「新聞」とCSJの「学会講演」には共通点が存在することが明らかになった。秋本(2016)では、BCCWJおよびCSJを用いて、「ちょっと・すこし」、「よ・ね」の出現頻度や使用特徴を考察し、両者の言語的差異が指摘されている。CSJはBCCWJよりも「チョット・スコシ」「ヨ・ネ」の出現比率が全体的に高い。また、BCCWJでは語の割合に大きな差が見られないのに対し、CSJでは「スコシ」より「チョット」、「ヨ」より「ネ」の割合が有意に高く、口語的特徴が顕著にあらわれる。一方で、ジャンル別に分析した場合、単純な書き言葉と話し言葉の二項対立とは異なる分類パターンが観察されることも示されている。陳(2022, 2024)は、書き言葉と話し言葉を組み合わせた調査が少ない背景を踏まえ、書き言葉コーパスであるBCCWJと話し言葉コーパスであるCEJCを用い、漢語動名詞の出現状況および日中同形語の使用実態を検証した。また、陳(2026)はBCCWJとCEJCを用いて、日本語における漢語動名詞の品詞用法と、ジャンルや文体との関連性を明らかにした。陳(2022, 2024, 2026)の一連の研究は、日本語教育において書き言葉と話し言葉を統合的に分析することの重要性や、ジャンル・場面が言語使用に与える影響の検証の必要性を示唆している。

以上のように、「は」と「が」の使い分けについては様々な研究が行われてきた。しかし、先行研究の多くは、文法的考察、単一コーパス分析、あるいは質問紙調査など、特定の方法に依拠しており、実際の言語使用の全体像総合的に捉えるには限界がある。具体的には、(1)書き言葉と話し言葉における多様な変種の違いを十分に明らかにしていない、(2)実際の日本語における「は」と「が」の使用頻度や使用傾向が十分に検証されていない、(3)「は」と「が」がどのような意味機能別で多く用いられているのかを体系的に解明していない、といった課題も残されていた。

3. リサーチデザイン

3.1 本研究の目的と研究設問

前述のように、本研究の目的は、外国人学習者のための「は」と「が」の指導方針を考えるための準備作業として、日本語の様々な言語変種における「は」と「が」の使用実態の解

明を目指すことである。この目的に沿い、BCCWJ¹(4 変種)と CEJC(3 変種)を資料として、以下の研究設問(research question : RQ)を設定する。

RQ1 「は」と「が」の頻度はいずれが優先するか？

RQ2 「は」と「が」の主要機能は相互にどのような関係にあるか？

3.2 使用するコーパス

BCCWJ(前川ほか, 2015)とは、現代日本語の書き言葉の実態を包括的に把握することを目的として構築された均衡コーパスである。書籍、雑誌、新聞、白書、ブログ、教科書、法律など、13 のジャンルから収集されたテキストが収録されており、総語数は約 1 億語に及ぶ。そのため、現代日本語の書き言葉の特徴を総合的に分析・研究するための基礎資料として広く利用されている。

CEJC(小磯ほか, 2023)とは、実際の日常会話の実態を把握することを目的として構築された会話コーパスである。多様な場面や話者属性を考慮して会話データを収集しており、会話場面は「雑談」「用談・相談」「会議・会合」「授業・レッスン」の 4 場面に分類されている。また、音声データおよび文字化データに加え、映像データも公開されている点に特徴がある。

本研究では、BCCWJ の主要 4 ジャンル(BCCWJ1 の書籍、BCCWJ2 の書籍、雑誌、知恵袋)と CEJC の 3 場面(雑談、用談・相談、会議・会合)を資料として、「は」と「が」の使用実態の調査を行うこととする。上述の 2 つのコーパスはコーパス検索アプリケーション「中納言」で無償に利用することができる。

3.3 分析手順

3.3.1 意味機能リストの作成

RQ2 では、機能別の頻度調査を行うが、その際必要になるのは、「は」と「が」の機能リストを作成することである。そこで、はじめに、国語辞典における「は」と「が」の語義分類を集約する。今回は、5 種の国語辞典(『新明解国語辞典 第六版 特装版』、『旺文社国語辞典 第十版』、『岩波国語辞典 第七版』、『三省堂国語辞典 第七版』、『明鏡国語辞典 第二版』)に含まれる係助詞「は」²と格助詞「が」の項目をすべて抽出し、重なりを除去した上で、2 語の意味機能を表 2 のように整理した。

¹ 現在、国立国語研究所では『現代日本語書き言葉均衡コーパス』(BCCWJ)の拡張として、『現代日本語書き言葉均衡コーパス 第 2 部』(BCCWJ2)を構築している。これに伴い、これまで『現代日本語書き言葉均衡コーパス』(BCCWJ)として公開してきたコーパスの名称を『現代日本語書き言葉均衡コーパス 第 1 部』(BCCWJ1)に変更する。本研究では、BCCWJ1 と BCCWJ2 を併用する。

(<https://clrd.ninjal.ac.jp/bccwj/>より、2026 年 7 月 15 日確認)

² 「副助詞」と分類される場合もある。

表 2 「は」と「が」の語義分類

助詞	機能用法	意味	例文
「は」	主題	個別具体的・一般抽象的な主題を表す。物事を他のものと区別して、取り立てて示すのに使う。	あの料亭 <u>は</u> 有名な <u>のか</u> 。 クジラ <u>は</u> 魚 <u>ではない</u> 。
	取り立て	話し手が話題にしたいこと(...について言えば)と、事柄の範囲や程度を取り立てて、それと限定したり、その範囲をことさらに強調したりする。	私 <u>は</u> 高橋と申します。[話題] 雨が降ると <u>は</u> 思っていない。[範囲]
	対比	他の物事と対比的に取り上げられることを表す。	声 <u>は</u> 聞こえるが、姿 <u>は</u> 見えない。 東京に <u>は</u> 行ったが、横浜に <u>は</u> 行っていない。
	絶対提示	「A は A だ」の形で、同じ語を「は」の前後で反復させて、内容を確認する気持ちを表す。どんな場合でも変わらないこと、独特であることも表す。	若く見える母も五十歳 <u>は</u> 五十歳 <u>なのだ</u> 。 規則 <u>は</u> 規則 <u>だ</u> 。 私 <u>は</u> 私で考えがある。
「が」	主体	動作や作用・性質・存在・受け手・使役者などの主語(主体)を表す。	雨 <u>が</u> 降る。[動作・作用] 空 <u>が</u> 青い。[性質] 太郎 <u>が</u> 先生に褒められた。 [受け手] 先生 <u>が</u> 花子に本を読ませた。[使役者]
	対象	感覚・感情・希望・好悪・可能(能力)などの対象を表す。	僕は足 <u>が</u> 痛い。[感覚] 故郷 <u>が</u> 懐かしい。[感情] 魚 <u>が</u> 食べたい。[希望] 映画 <u>が</u> 好きだ。[好悪]
	排他	他のものを除外して、「こうであるのはこれだけであり、他のものは該当しない」と限定・特定する働きのことを表す。	でも冬 <u>が</u> 一番出るね。 応募の作品の中で、この作品 <u>が</u> 一番いい。
	非難	文末にののしる気持ちを表す。	(この)おろかもの <u>が</u> ...
	特別化	「A が A」の形で着目する A の性質に注意を喚起し、改めて A だと確認する気持ちで言うのに使う。一般には、同じことばの間にはさんで、対象が特別のものごとであることを表す。	場合 <u>が</u> 場合だから。 事情 <u>が</u> 事情なので肩身が狭い。

3. 3. 2 コーパスからの用例抽出

本研究では、前述の 3.2 で示した通り、BCCWJ の主要 4 ジャンルと CEJC の 3 場面、

計 7 変種を対象とする。「中納言」 インタフェースを用い、検索単位を短単位、単語タイプを語彙素、検索対象を「は」と「が」、品詞の小分類を「は」は「助詞-係助詞」、「が」は「助詞-格助詞」と設定する。このうえ、7 変種における「は」と「が」の用例については、検索結果からランダムに抽出された各 500 例を Excel 形式で保存する。さらに、各変種から「は」「が」の用例を無作為に 100 ずつ、計 1,400 例を収集する。

3. 3. 3 RQ ごとの手法

RQ1(頻度関係の変化)では、「中納言」 インタフェースを用い、先述した検索方法に基づき、BCCWJ と CEJC の計 7 変種を対象として、係助詞「は」と格助詞「が」の出現頻度を調査する。各変種間ではコーパスの規模が異なるため、単純な出現頻度の比較は適切ではない。そこで、石川(2021)に基づき、コーパスサイズの違いによる影響を補正するため、100 万語当たりの出現頻度(Per Million Words, PMW)を算出し、調整頻度として比較を行う。さらに、「は」と「が」の PMW の合計値に占める「は」の割合を算出し、変種間における両助詞の使用傾向を比較・分析する。

RQ2(主要機能の変化)では、まず、各変種から無作為に抽出した「は」と「が」の用例計 1,400 例を分析対象とする。次に、上記の表 1 に示した意味機能分類に基づき、Excel を用いて各用例を手作業で意味機能に分類した上で、意味機能ごとの出現比率を算出する。その後、フリーの分析ソフト HAD18.0(清水, 2023)を用いて階層的クラスター分析や対応分析を実施し、「は」と「が」の意味機能の分布傾向を可視化・比較することにより、両助詞の主要機能の変化の特徴を明らかにする。

4. 結果と考察

4.1 7 変種における「は」と「が」の頻度

書き言葉と話し言葉を含む 7 変種において、「は」と「が」の頻度的な優先性を調査すべく、「は」と「が」の総頻度に占める「は」の占有率を調査したところ、表 3 の結果を得た。

表 3 各変種における PMW と「は」の占有率

	言語変種	は(PWM)	が(PWM)	はの占有率
BCCWJ	書籍(BCCWJ1)	29048.36714	20424.94807	58.72%
	書籍(BCCWJ2)	27296.42617	20369.48299	57.27%
	教科書	22731.61482	18761.48321	54.78%
	知恵袋	25708.2099	21182.3129	54.83%
CEJC	雑談	14144.35586	15319.39023	48.01%
	用談・相談	16159.89841	15518.29783	51.01%
	会議・会合	18350.14307	16393.82141	52.82%

全体的に見ると、書き言葉ではすべてのジャンルで「は」の割合が 50%を超えており、「は」が「が」より優先的に使用される傾向が確認された。特に、「書籍(BCCWJ1)」(58.72%)が最も高く、次いで「書籍(BCCWJ2)」(57.27%)、「知恵袋」(54.83%)、「教科書」(54.78%)

の順となっている。ジャンル間の差は約 4 ポイントと比較的小さく、書き言葉ではジャンルの違いにかかわらず、「は」が一貫して優勢であることが確認できた。

一方、話し言葉では場面によって異なる傾向が見られた。「用談・相談」(51.01%)および「会議・会合」(52.82%)では、「は」の割合が 50%を上回っており、書き言葉と同様に「は」が優勢であった。しかし、「雑談」では 48.01%と 50%を下回り、「が」の割合が 51.99%となることから、「が」の使用が「は」を上回る唯一の場面であることが分かる。

この結果から、日本語の言語的特徴の観点から言えば、これは日本語において主格明示よりも取り立て機能が優先されている可能性を示唆している。書き言葉では情報を主題化しながら論理的に文章を展開する必要があるため、「は」が安定して高い割合で使用されるのに対し、話し言葉では会話の目的や場面によって助詞の選択が変化することが示唆される。特に、自由な対話を特徴とする「雑談」では、新情報や話題の導入が頻繁に行われるため、「が」の使用が相対的に増加したと考えられる。一方、「用談・相談」や「会議・会合」では、既知の話題について継続的に議論を進める場面が多く、話題を維持する機能を持つ「は」が優先される傾向があると考えられる。

以上のことから、全体として「は」の優先が確認された一方、「は」と「が」の選択は、単に書き言葉・話し言葉という言語変種の違いだけでなく、コミュニケーションの目的や談話構造にも大きく影響されることも明らかとなった。特に、「雑談」のみが「が」優勢となった点は、話し言葉の中でも場面によって助詞選択の傾向が異なることを示す特徴的な結果である。

4.2 「は」と「が」の主要機能の相互関係

(言語変種【7 変種】)×(2 語の主要機能【9 種】)の頻度表にクラスター分析と対応分析を適用し、2 語の主要機能の相互関係を観察したところ、以下の結果を得た。なお、対応分析で得られた第 1 次元寄与率は 47.1%、第 2 次元寄与率は 27.0%であり、累積寄与率は 74.1%であり、2 つの次元によって分散の過半が説明されていると考えられる。

表 4 「は」と「が」の主要機能

	は				が				
	主題	取り立て	対比	絶対提示	主体	対象	排他	非難	特別化
書籍 (BCCWJ1)	67	19	14	0	69	18	11	1	1
書籍 (BCCWJ2)	67	21	12	0	76	17	7	0	0
教科書	61	26	13	0	75	20	5	0	0
知恵袋	58	26	16	0	72	16	11	0	1
雑談	76	10	13	1	71	21	7	0	1
用談・相談	72	22	5	1	77	21	1	0	1
会議・会合	70	20	8	2	77	16	7	0	0

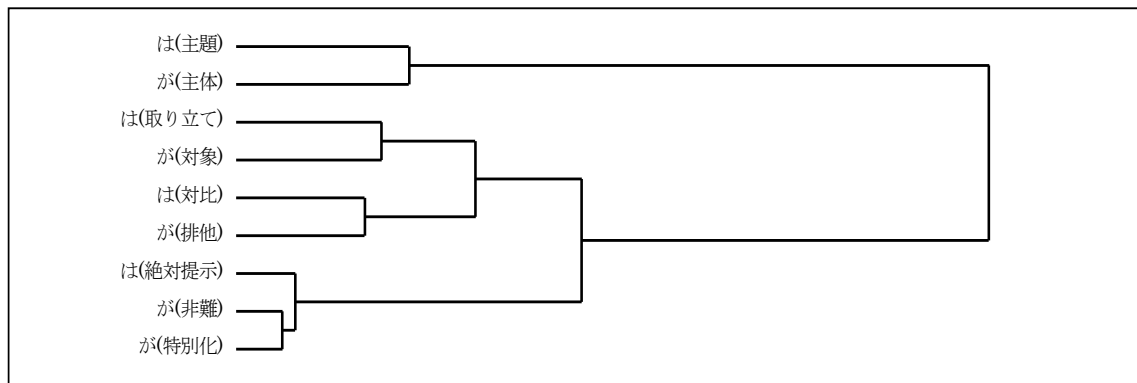


図1 「は」と「が」の主要機能の階層関係

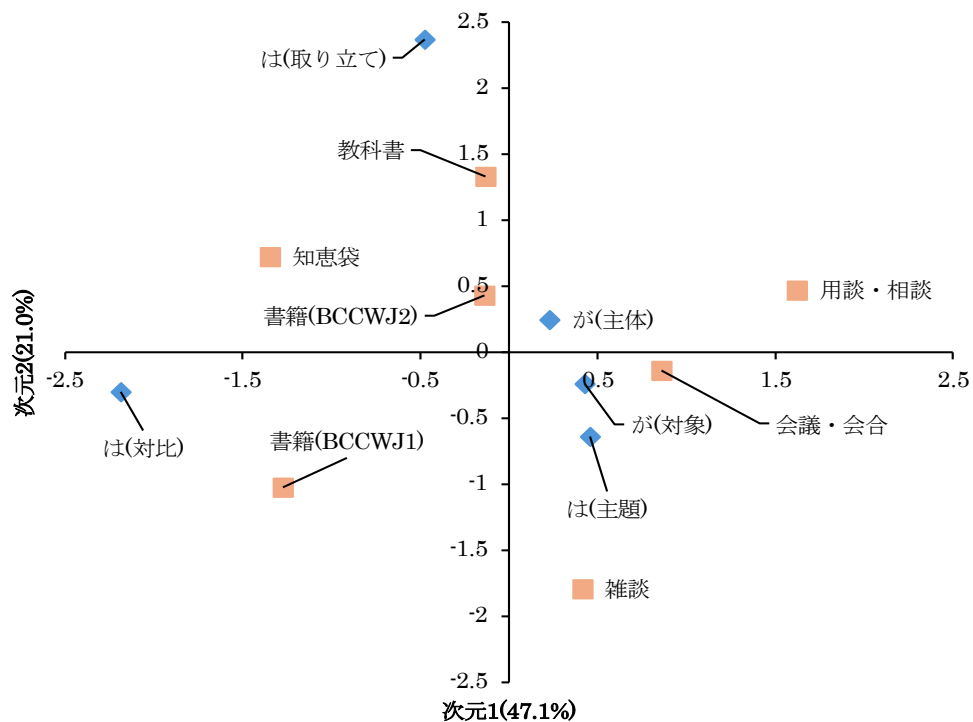


図2 「は」と「が」の主要頻度の対応関係

図1より、「は」と「が」の意味機能は大きく三つのグループに分類されることが確認できた。まず、「主題(は)」と「主体(が)」は最も近い距離で結合しており、さらに他の機能群とは最も遅い段階で統合されている。このことから、両者は他の意味機能とは異なる独立性の高いグループを形成していることが分かった。両機能はいずれも文の基本的な構成要素を担う機能であり、「は」が主題を提示し、「が」が述語の主体を示すという、最も基本的な用法であることを反映していると考えられる。

次に、「取り立て(は)」「対象(が)」「対比(は)」「排他(が)」は一つのクラスターを形成している。このうち、「取り立て(は)」と「対象(が)」がまず結合し、「対比(は)」と「排他(が)」が別の小クラスターを形成した後、両者が統合されている。この結果は、「は」の取り立て

機能と「が」の対象表示機能、および「は」の対比機能と「が」の排他機能が、それぞれ比較的類似した分布傾向を示していることを意味している。これらはいずれも、情報の焦点化や選択性、対照性といった談話的な機能に関わる点で共通性を有していると解釈できた。

さらに、「絶対提示(は)」「非難(が)」「特別化(が)」は別のクラスターを形成している。「非難(が)」と「特別化(が)」は最も近い距離で結合しており、両機能の分布傾向が極めて類似していることが確認された。その後、「絶対提示(は)」がこれらと統合され、一つのクラスターを構成している。この結果から、これら三つの機能はいずれも出現頻度が極めて低く、特殊な談話環境や文脈に依存する用法であるという共通した特徴を有する可能性が示唆される。もっとも、本研究で分析対象とした全 1,400 例のうち、「絶対提示(は)」「非難(が)」「特別化(が)」に該当する事例はわずか 9 例にとどまっており、サンプル数が極めて少ない。そのため、本クラスターの形成には偶然性が含まれている可能性も否定できず、その妥当性については今後、分析対象を拡大した上で再検証する必要がある。したがって、本研究ではこのクラスターの解釈については慎重な立場を取り、詳細な考察の対象とはしない。

全体として、「主題(は)」と「主体(が)」は他の意味機能とは明確に区別される一方で、それ以外の意味機能は談話機能や情報構造との関連性に応じて複数のグループを形成していることが明らかとなった。特に、「は」と「が」の対応する機能が必ずしも同一のクラスターに属しているわけではなく、「取り立てー対象」や「対比ー排他」のように、異なる助詞間で類似性が認められた点は興味深い。このことは、「は」と「が」の意味機能は単純に助詞ごとに分類されるのではなく、情報構造や談話機能という観点から相互に関連しながら体系を形成していることを示唆している。

続いて、図 2 に関して、はじめに、第 1 次元を見ると、左側には「教科書」「書籍(BCCWJ1)」「書籍(BCCWJ2)」「知恵袋」といった書き言葉のジャンルが配置され、右側には「会議・会合」「用談・相談」「雑談」といった話し言葉の場面が配置されている。また、「取り立て(は)」と「対比(は)」は書き言葉ジャンルに近接している一方、「主体(が)」「対象(が)」「主題(は)」は話し言葉の場面の近くに位置している。このことから、第 1 次元は書き言葉と話し言葉を区別する軸であると解釈できる。すなわち、相対的に見ると、書き言葉では「は」の取り立てや対比の機能が相対的に多く用いられるのに対し、話し言葉では「は」の主題提示や「が」の主体・対象を表す機能が多く用いられる傾向が示された。

次に、第 2 次元では、「取り立て(は)」と「教科書」が図の上方に位置し、「主題(は)」「対象(は)」「雑談」が下方に配置されている。この配置から、第 2 次元は談話の形式性や情報提示のあり方を反映する軸であると考えられる。すなわち、上方には説明的・客観的な性格を持つ変種が位置し、下方には会話的・相互作用的な性格を持つ変種が位置する傾向が認められる。

さらに、各言語変種と助詞機能との対応を見ると、「教科書」は「取り立て(は)」に最も近接しており、説明対象を提示し、その内容について記述・解説を行うという教科書の特徴を反映していると考えられる。一方、「書籍(BCCWJ1)」は「対比(は)」に近く、人物や事象を対照的に描写する表現との関連性が示唆される。また、「書籍(BCCWJ2)」と「知恵袋」は原点付近に位置しており、特定の助詞機能との強い対応は認められず、複数の機能が比較的均等に用いられていることがうかがえる。

話し言葉の場面に着目すると、「用談・相談」は「主体(が)」と比較的近い位置にあり、主体を明示する表現との関連が強いことが示されている。また、「会議・会合」は「対象(が)」と「主題(は)」に近接しており、議論の対象を示す表現と、談話の主題を提示する表現がと

もに多く用いられる傾向が認められる。さらに、「雑談」は第2次元の下方に独立して位置しており、他の変種と比較して口語的・対話的な特徴が強いことを示している。

以上の結果から、「は」と「が」の主要機能は言語変種によって明確な偏りを示しており、2語の機能を相対的に見ると、書き言葉では「は」の「取り立て」・「対比」機能が中心となる一方、話し言葉では「は」の「主題」や「が」の「主体」・「対象」を表す機能が中心となることが明らかとなった。このことは、「は」と「が」の使い分けが文法的要因だけではなく、言語変種や談話の性質、さらには情報構造とも密接に関連していることを示唆している。

5. まとめ

以上、本研究では、BCCWJとCEJCを資料として、計7変種のデータを用いて、現代日本語の書き言葉と話し言葉における「は」と「が」の出現状況について計量的検証を行ってきた。以下、得られた結果をまとめる。

まず、RQ1(頻度関係の変化)については、7言語変種中6変種で「は」が優先され、日本語においては主格明示よりも取り立て機能が優先されている可能性が示された。媒体別に見ると、書き言葉では「は」、話し言葉では「が」との親和性が相対的に高いことが明らかとなった。ただし、話し言葉の中では「雑談」においてのみ順位の逆転(「は」<「が」)が生じることが判明した。これは、「雑談」が同一の時空間を共有する文脈であり、知覚した現場の状況や突発的な事象を共有する「現象文」が多用されるためであると解釈できる。一方、書き言葉では読者と時空間を共有しないため、言語表現によって主題や背景の土台を丁寧に形成する必要があり、「は」の選択確率が高まると考察された。

次に、RQ2(主要機能の変化)については、階層的クラスター分析と対応分析を通して、「は」と「が」の主要機能と変種との対応関係を明らかにした。階層的クラスター分析では、「主題(は)」と「主体(が)」が独立したクラスターを形成し、両者が文の基本構造を担う中核的機能であることが示された。一方、「取り立て(は)」と「対象(が)」、「対比(は)」、「排他(が)」は、それぞれ類似した分布傾向を示し、助詞の違いを超えて情報構造や談話機能との関連性によってまとまりを形成していることが確認された。また、対応分析では、書き言葉は「は」の「取り立て」や「対比」機能との対応が強く、話し言葉は「は」の「主題」機能と「が」の「主体」や「対象」機能との対応が強いことが示された。すなわち、「は」と「が」の使い分けは単なる文法的対立ではなく、言語変種や談話の性質、情報構造と密接に関連して変化することが明らかとなった。

本研究で得られた知見は、外国人学習者のための「は」と「が」の指導の指針に対して、いくつかの教育的示唆を与えられる。従来の日本語教育では、「は」は主題、「が」は主語という文法的な対立を中心に説明されることが多い。しかしながら、本研究の結果から、「は」と「が」の使い分けは文法的な要因だけでなく、書き言葉・話し言葉という媒体の違いや、言語変種、さらには情報構造や談話機能とも密接に関わっていることが解明された。そのため、助詞の指導においては、単に文法規則や例文を提示するだけでなく、実際の言語コーパスに基づく使用傾向を取り入れ、言語変種ごとの特徴を意識した指導を行うことが重要である。

今後の課題として、本研究ではBCCWJとCEJCのみを調査対象としたが、これら二つのコーパスのみでは、現代日本語の書き言葉と話し言葉の実態を十分に反映しているとは言えない。そのため、本研究で得られた知見については、他の代表的な大規模コーパスを用いて再検証する必要がある。また、BCCWJとCEJCの変種から抽出した用例数も十分と

は言えず、分析結果には一定の偶然性が含まれる可能性がある。今後は、分析対象となるデータ量をさらに拡充し、より深い検証を行うことで、本研究で得られた結論の妥当性や一般性をより高めていくことが求められる。

文 献

- 秋本瞳(2016)「コーパスにみる話しことばと書きことばの連続性—BCCWJ と CSJ におけるチョット/スコシ, ヨ/ネの出現頻度の比較を通じて—」『言語と文明』14, 21-41.
- 庵功雄(2016)「「産出のための文法」から見た「は」と「が」」庵功雄・佐藤琢三・中俣尚己編『日本語文法研究のフロンティア』くろしお出版.
- 庵功雄(2018)「「は」と「が」の新しい捉え方についての一考察:「は」と「が」はこんなに簡単だった!」『一橋日本語教育研究』6, 1-10.
- 庵功雄(2019)「コーパスを用いて「は」と「が」に関する三上説を検証する試み」『言語資源活用ワークショップ発表論文集』199-204.
- 庵功雄(2021)「「は」と「が」の使い分けのあり方を定量的に確認する試み」『言語資源活用ワークショップ発表論文集』6, 198-203.
- 石川慎一郎(2013)「テキスト関連属性と助詞選択: 計量的アプローチに基づく探索的研究—主語・主題を導く「は」と「が」をめぐる—」『第4回コーパス日本語学ワークショップ予稿集』213-222.
- 石川慎一郎(2021)『ベーシック コーパス言語学 第2版』くろしお出版.
- 奥野由紀子, ディアンニ・リスダ(2015)「「話す」課題と「書く」課題に見られる中間言語変異性」『国立国語研究所論集』9, 121-134.
- 北原保雄(2012)『明鏡国語辞典 第二版』大修館書店.
- 見坊豪紀, 市川孝, 飛田良文, 山崎誠, 飯間浩明, 塩田雄大(2013)『三省堂国語辞典 第七版』三省堂.
- ゲオルギエバ・ベロニカ(2012)「「は」と「が」の使い分けは二者択一か—日本語母語話者と日本語学習者による空所補充の比較から」『早稲田日本語教育学』12, 23-42.
- 小磯花絵・天谷晴香・居關友里子・臼田泰如・柏野和佳子・川端良子・田中弥生・伝康晴・西川賢哉・渡邊友香(2023)「『日本語日常会話コーパス』設計と構築」『国立国語研究所論集』24, 153-168.
- 小口悠紀子(2017)「上級日本語学習者の談話における「は」と「が」の知識と運用——未出か既出かによる使い分けに着目して——」『日本語教育』166, 77-92.
- 国際交流基金(2026)「海外の日本語教育の現状: 2024 年度 海外日本語教育機関調査より」<https://www.jpf.go.jp/j/project/japanese/survey/result/dl/2024/all.pdf> 2026 年 7 月 27 日閲覧
- 国立国語研究所(2025)『現代日本語書き言葉均衡コーパス』(バージョン 2021.03, 中納言バージョン 2.7.3, 分類語彙表情報 2025.03)<https://clrd.ninjal.ac.jp/bccwj/>(2026 年 7 月 15 日確認)
- 国立国語研究所(2026)『現代日本語書き言葉均衡コーパス 第2部』(バージョン 2026.03, 中納言バージョン 2.7.3) <https://www2.ninjal.ac.jp/BCCWJ2/>(2026 年 7 月 15 日確認)
- 清水裕士(2016)「フリーの統計分析ソフト HAD: 機能の紹介と統計学習・教育, 研究実践における利用方法の提案」『メディア・情報・コミュニケーション研究』1, 59-73.
- 朱新蕊(2024).「中国日语学习者日语助词“は”与“が”的习得研究(日本語訳:中国人日本語学習

- 者における助詞「は」と「が」の習得に関する研究) 黒竜江大学修士論文.
- 杉村泰(2010)「コーパスから見た中国人日本語学習者の格助詞に関する問題点について」『言語文化研究叢書』9, 137-152.
- 高見澤孟・ハント蔭山裕子・池田悠子・伊藤博文・宇佐美まゆみ・西川寿美(2004)『新・はじめての日本語教育 1: 日本語教育の基礎知識』アスク出版.
- 滝島雅子(2020)「話し言葉と書き言葉における敬語名詞の語彙比較 CEJC と BCCWJ のデータを用いて」『計量国語学』32(6), 315-330.
- 陳迪(2022)「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」『統計数理研究所共同研究リポート』456, 21-41.
- 陳迪(2024)「中国語を母語とする日本語学習者による日中同形語使用の問題点: 多様なコーパスデータを用いた調査」『Learner Corpus Studies in Asia and the World』6, 147-168.
- 陳迪(2026)漢語動名詞の品詞用法に見られるジャンル差と文体特性: 頻度分析および特徴語分析に基づいて」『統計数理研究所共同研究リポート』484, 1-18.
- 中西久実子(2024)「日本語学習者の従属節を含む文の「は・が」の誤用の実態: 事例の長文を使用した誤用の実態調査から」『国際言語文化』8, 31-46.
- 日本語記述文法研究会(2009)『現代日本語文法 5: 第 9 部とりたて; 第 10 部主題』くろしお出版.
- 西尾実, 岩淵悦太郎, 水谷静夫(2009)『岩波国語辞典 第七版』岩波書店.
- 野田尚史(1991)『はじめての人の日本語文法』くろしお出版.
- 野田尚史(1996)『新日本語文法選書 1: 「は」と「が」』くろしお出版.
- 彭広陸, 守屋三千代(2022)『総合日語 第一冊 第三版(日本語訳: 総合日語 第一冊 第三版)』北京大学出版社.
- 前川喜久雄(監修), 山崎誠(編)(2014)『書き言葉コーパス—設計と構築—』, 講座日本語コーパス 2. 朝倉書店.
- 松村明, 山口明穂, 和田利政(2005)『旺文社国語辞典 第十版』旺文社.
- 丸山直子(2015)「助詞の使用実態—BCCWJ・CSJ にみる分布—」『第 8 回コーパス日本語学ワークショップ予稿集』179-188.
- 山田忠雄, 柴田武, 酒井憲二, 倉持保男, 山田明雄(2004)『新明解国語辞典 第六版 特装版』三省堂.
- 李在鎬(2020)「日本語教育学の課題に対して計量分析は何ができるか」『計量国語学』32(7), 372-386.
- 劉利国, 宮偉(2019)『新经典日本语基础教程 第 1 册 第 2 版(日本語訳: 新经典日本語基礎教程 第 1 册 第 2 版)』外語教学与研究出版社.

関連 URL

コーパス検索アプリケーション『中納言』

<https://chunagon.ninjal.ac.jp/>