



コーパスに基づく漢語動名詞の連体修飾後続要素についての調査

陳, 迪

(Citation)

統計数理研究所共同研究レポート, 469:1-10

(Issue Date)

2024-03-25

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/0100487702>

(URL)

<https://hdl.handle.net/20.500.14094/0100487702>



コーパスに基づく漢語動名詞の連体修飾後続要素についての調査

陳 迪 (神戸大学大学院生)

A Corpus-based Study of Postnominal Elements of Sino-Japanese Verbal Nouns

CHEN, Di (Kobe University, Graduate Student)

概要

本研究では、日本語学習者が理解できる漢語動名詞の連体修飾後続要素の使い分け基準を作成するため、「スル」・「ノ」・「的」・「的ナ」の許容度を調査した上で、連体修飾後続要素に影響を及ぼす要因を探索した。その結果、(1)「スル」は最も典型的な連体修飾後続要素であること、(2)漢語動名詞が法律、白書、国会会議録などのような政治・社会に関する公的な文書に出現する場合には「スル」を選択する確率が高いこと、が確認された。

キーワード

コーパス、漢語、動名詞、連体修飾後続要素

1. はじめに

日本語には、「準備」「利用」などのような、名詞自体に動詞的意味を有し、「する」を付加することで動詞としても使われる漢語が存在する。このような漢語は漢語動名詞 (Sino-Japanese verbal nouns : 以下、漢語 VN) と呼ばれている (小林, 2004; 桑原, 2017)。論理的に言えば、漢語 VN は名詞と動詞の両面性を持っているため、「漢語 VN ノ N」と「漢語 VN スル N」の 2 種類の構造で名詞を修飾し、連体修飾語となることができる。以下は、「現代日本語書き言葉均衡コーパス」 (Balanced Corpus of Contemporary Written Japanese: BCCWJ) からの用例である。

- (1) だから、あらかじめ準備する工具や部品などには気を遣う。(PM35_00034, 出版・雑誌)
- (2) または外国為替準備の減少によってまかなわれる (LBo3_00139, 図書館・書籍)
- (3) 知恵袋を利用する人は子供からお年寄りまでいますからね。(OC14_09796, 特定目的・知恵袋)
- (4) ローンのしくみを理解して、賢い利用の方法をきちんと覚えることが大切です。(PB25_00153, 出版・書籍)

上記のうち、(1)と(3)では「漢語 VN スル N」が選択されているが、(2)と(4)では「漢語 VN ノ N」

が選択されている。さらに、BCCWJには下記のような用例も存在する。

- (5) 世界貿易の安定的発展のために早い機会に規律の確立が行われることが望ましい。
(OW1X_00074, 特定目的・白書)
- (6) 我が国が世界各国との安定的な関係を築くうえでもきわめて重要である。(OW2X_00169, 特定目的・白書)

「安定」は漢語 VN であるが、「的」および「的ナ」が後続して名詞を修飾し、形容詞的性質を持つ場合もあることがわかった。

すなわち、漢語 VN は連体修飾語として使われる際に、少なくとも「スル」・「ノ」・「的」・「的ナ」の4つの連体修飾後続要素との共起が可能であることが確認された。しかしながら、漢語 VN は常に4種の後続要素が取れるとは限らない。たとえば、石(2021)は、漢語 VN の場合、「VNノN」の成立には、VNがNの性質や特徴を規定しているという制約条件が存在することと指摘した。

なお、中国語を母語とする日本語学習者(L1 Chinese Learners of Japanese : 以下、CLJ)にとって、漢語 VN の連体修飾用法の習得は必ずしも容易ではない。以下は「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)からのCLJと日本人母語話者(Japanese Native Speaker : 以下、JNS)の用例である。

- (7) 公園に着いたあと、朝準備の食べ物を食べるつもり時… (CCT22-SW1)
- (8) その中に、せっかく準備の食べ物は全部食べられました。(CCT22-SW1)
- (9) お昼の時間になり、朝二人で準備したサンドイッチを食べようと… (JJJ05-SW1)

上記の3つの用例から見ると、同じ意味であっても、CLJは「ノ」を選択している(例(7)と(8))のに対し、JNSは「スル」を選択している(例(9))。実際に、BCCWJを調査すると、「準備」が「漢語 VNノN」の構造として使われる例は少なくないが、Nには、事柄の構成部分(「段階」「一環」など)・状況の変化(「急増」「減少」など)などを表す名詞が多いのに対して、具体的な事物(「食べ物」「資料」など)がほとんどない。すなわち、「準備」は「漢語 VNノN」の構造で連体修飾語として使われる場合、修飾される名詞の性質や特徴を規定することができない。しかし、中国語は語形変化がないため、名詞・動詞・形容詞を問わず、連体修飾の場合には一括して“的”が使われる。CLJの漢語動名詞の連体修飾用法には誤用が多いのは、母語干渉によるのではないかと考えられる。

上記をふまえると、漢語 VN では、語によって、異なる連体修飾後続要素を選択する可能性が存在し、CLJは母語からの干渉によって誤用を生じやすいことがわかった。したがって、本研究では、コーパスデータを用い、漢語動名詞の連体修飾要素である「スル」・「ノ」・「的」・「的ナ」の許容度を調査した上で、連体修飾後続要素に影響を及ぼす要因を探索する

ことで、CLJ が理解できる連体修飾後続要素の使い分けの基準を明確にすることを旨とする。

2. 先行研究

日本語漢語の連体修飾用法については様々な研究が行われてきたが、ここではコーパスデータに基づく事例調査に絞って概観する。

石(2021)は、寺村(1993)が提唱された連体節構造に関する概念である「内の関係」と「外の関係」を踏まえて、新たな視点から漢語動名詞の連体用法「VN ノ N」の成立条件を考察した。具体的には、BCCWJ を用いて、「(漢語) VN ノ N」の用例を収集し、「内の関係」と「外の関係」の定義に従って整理・分類した。その結果、(1) 外の関係「VN ノ N」は全体の 96.6% を占めており、一般的に成立可能である、(2) 内の関係「VN ノ N」は全体の 3.4% を占めており、一般に成立不可であるが、特殊な場合にも成立可能である、ことが明らかになった。さらに各種の漢語動名詞の「VN ノ N」について考察すると、(a) 外の関係「VN ノ N」の場合、内容型・基準型・随伴物型の 3 種類に分けられ、(b) 内の関係「VN ノ N」の特殊な用法の場合、VN は主名詞の N の特徴や属性を表し、つまり、形容詞の役割を持っている、のことを指摘した。なお、50 名の CLJ を対象に、「VN {スル/シタ、ノ} N」を選択させる問題(30 問)を用意して、CLJ の「VN ノ N」の使用意識を調査した。その結果、(ア) 「スル/シタ」しか使えない場合に「ノ」を選択した CLJ の非正答率は 67.6% であった、(イ) 「ノ」と「スル/シタ」両者ともに成立するが、どちらか一方しか選択しなかった非正答率は 62% であった、ことがわかった。このような結果は、多くの CLJ は漢語動名詞の連体修飾用法を使い分けできないことが示唆された。

鄧(2022)は、BCCWJ を用い、漢語の連体修飾用法について調査を行った。まず、各連体修飾パターンの頻度に基づいて漢語 100 語を分類した。漢語動名詞のみに注目すると、(1) 「発生」「参加」「対応」「利用」「使用」「実施」などの 5 語は「シタ」型、(2) 「意味」「関係」「生活」「研究」などの 20 語は「ノ」型、(3) 「説明」「確認」「開発」「理解」などの 4 語は「ノ」・「シタ」型、さらに、(4) 「代表」は「的ナ」型、(5) 「計画」「自動」は「ノ」・「的ナ」型、(6) 「決定」は「ノ」・「シタ」・「的ナ」型、であることがわかった。これによって、漢語動名詞は連体修飾語として使用する際に、少なくとも 6 種のパターンがあることが明らかになった。

次に、鄧(2022)は「形状詞可能型」の漢語を、「ナ型」・「ナ寄り型」・「両用型」・「ノ寄り型」・「ノ」型に分類し、各型に属するものはそれぞれどのような言語特性を持っているかについて調査した。その結果、漢語の言語特性による「ナ」と「ノ」の選択モデルが得られ、関係概念を表す場合「ノ」を後続することが確認された。

本研究では、上記の石(2021)と鄧(2022)の研究手法を踏襲して、漢語動名詞の連体修飾用法について調査を行う。

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、本研究は日本語学習者が理解できる漢語 VN の連体修飾後続要素の使い分け基準を作成するため、「スル」「ノ」「的」「的ナ」の許容度を調査した上で、連体修飾後続要素に影響を及ぼす要因を探索する。この目的に従い、以下の 2 つのリサーチクエスチョンを設定する。

RQ1 各種の漢語 VN をまとめて見た場合、後続要素として「スル」「ノ」「的」「的ナ」のいずれが多いか？また、後続要素によってどのように分類するか？

RQ2 主要漢語 VN の「スル」後続に影響を及ぼす要因は何か？

3.2 データ

3.2.1 調査対象語

本研究では、陳 (2023) が抽出された現代日本語で一般的に多用される重要漢語 VN200 語の中から、上位 50 語をサンプル語として選んで調査を行う。陳 (2023) では、BCCWJ の内容別 65 変種および日本語日常会話コーパス (CEJC) の場面別 7 変種、あわせて 72 変種の言語資料で頻度調査を行い、平均頻度とレンジを手掛かりとして抽出した。上位 50 語の内訳を表 1 に示す。

表 1 重要漢語動名詞 (上位 50 語、重要度順)

利用	実施	使用	対応	紹介	確認	増加	発生	表示	作成	説明	参加	推進
発表	存在	理解	表現	期待	活用	注意	設定	設置	開催	減少	変化	購入
提供	比較	検討	結婚	選択	指定	展開	判断	入力	用意	検索	質問	確保
工夫	実現	評価	計算	開始	変更	発見	登録	注目	拡大	規定		

3.2.2 使用するコーパス

本研究では、日本語コーパスとして前述の BCCWJ と CEJC を、中国語コーパスとして Lancaster Corpus of Mandarin Chinese (LCMC) (McEnery & Xiao 2004) を使用する。

BCCWJ は国立国語研究所が中心となって構築された日本国内ではじめての大規模均衡コーパスである。新聞・雑誌・書籍・法律・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更に詳細な下位区分がある。

CEJC は国立国語研究所が 2016 年から構築している大規模話し言葉コーパスである。CEJC は様々なタイプの日本語日常会話およそ 200 時間をバランスよく収録した。会話の形式によって雑談、用談・相談、会議・会合、授業・レッスンの 4 種のジャンルに区分されている。

上述した BCCWJ と CEJC のデータはコーパス検索アプリケーション中納言で利用することができる。

LCMC は Richard Xiao が作成した現代中国語均衡コーパスである。該当コーパスは FLOB (Freiburg-LOB Corpus of British English) の構築モデルに従って作成され、デー

体量はおよそ 100 万語である。このコーパスは、中国語の書き言葉を想定母集団にし、1989 年から 1993 年まで出版された中国大陸の出版物に限ることを決めた。その後、15 種のカテゴリーごとに現実母集団を確定し、無作為抽出法によってサンプリングを決めた。LCMC は、CQPweb at Lancaster (Corpus Query Processor) というオンラインのシステムで無償に利用することができる。

3.3 手法

RQ1 では、重要漢語 VN 上位 50 語に対して、BCCWJ の全 13 ジャンルにおいて「漢語 VN-スル/シタ-名詞」、「漢語 VN -ノ-名詞」、「漢語 VN-的-名詞」、「漢語 VN-的-の-名詞」で検索し、得られた各後続要素共起の出現頻度を調査し、その比率を計算する。

RQ2 では、「スル」の共起率を目的変数として、目的変数にどのような要因が影響を及ぼしているかについて重回帰分析を用いて検討する。「スル」共起率の計算式は以下の通りである。

$$\text{「スル」共起率} = \text{「スル」共起の出現頻度} / \text{該当漢語 VN の出現頻度}$$

なお、漢語 VN の「スル」共起率に関連するものとして様々な要因があるが、ここでは、表 2 に示した要因を説明変数として挙げる。

表 2 重回帰分析の説明変数

変数性質	説明変数	内訳
量的変数	書き言葉頻度	個々の漢語 VN について BCCWJ での出現頻度を調査し、100 万語あたりの調整頻度 (PMW) に調整する。
	話し言葉頻度	個々の漢語 VN について CEJC での出現頻度を調査し、PMW に調整する。
	中国語頻度	中国語書き言葉コーパス LCMC での出現頻度を調査し、PMW に調整する。
	メディアジャンルの出現率	個々の漢語 VN について、BCCWJ 内の 13 ジャンル中、メディア性が強い「新聞」「雑誌」という 2 ジャンルにおける出現率を調査する。
	文書ジャンルの出現率	個々の漢語 VN について、BCCWJ 内の 13 ジャンル中、政治・社会に関する公的な文書である「白書」「法律」「国会会議録」という 3 ジャンルにおける出現率を調査する。
	カジュアルジャンルの出現率	個々の漢語 VN について、BCCWJ 内の 13 ジャンル中、カジュアル性が強い「ブログ」「知恵袋」という 2 ジャンルにおける出現率を調査する。

	画数	個々の漢語 VN について、 https://kanjijoho.com/ というウェブサイトを用いて画数を調査する。
	語義数	個々の漢語 VN について、『新明解』『大辞林』『岩波国語辞典』における語義数を調査して、平均値を求める。
ダミー 変数	意味分類	個々の漢語 VN について、『分類語彙表』に準じて語義分類を調査する。なお、該当する意味分類は「抽象的關係」「精神および行為」「自然現象」「人間活動の主体」という 4 種類であることため、4 カテゴリーごとに合致の有無をコード化する。
	自他性	個々の漢語 VN について、上述の 3 つの辞書における自他性を調査し、2 種以上に出現する品詞性を用いる。自他性については、「自動詞」「他動詞」「自他両用」の 3 分類を用いる。ここでは、「自動詞」と「他動詞」の 2 つのカテゴリーごとに合致の有無をコード化する。
	語構成要素の品詞性	個々の漢語 VN について、漢語を構成する 2 つの漢字の品詞性を調査する。なお、該当する品詞性パターンは、「動詞-動詞 (VV)」「副詞-動詞 (MV)」「動詞-形容詞 (VA)」「名詞-動詞 (NV)」「動詞-名詞 (VN)」「名詞-名詞 (NN)」という 6 種類があるため、6 カテゴリーごとに合致の有無をコード化する。
	品詞性	個々の漢語 VN について品詞用法を陳 (2023) の調査結果に基づき、「名詞的名詞」「動詞的名詞」「名詞的動詞」「動詞的動詞」の 4 種類に分類し、4 つのカテゴリーごとに合致の有無をコード化する。
	中国語品詞性	個々の漢語 VN について、『現代漢語詞典』(第 7 版)に基づき、中国語における対応する語の品詞性を調査する。なお、該当する中国語品詞性は「名詞」「動詞」「形容詞」「副詞」「介詞(動作の時間・場所・原因・対象などの意味関係を表す品詞)」の 5 種類に分類し、5 つのカテゴリーごとに合致の有無をコード化する。
	出現年代	個々の漢語 VN について『日本国語大辞典』に基づき、テキストに出現した最初の年代を調査する。例えば、「準備」の場合、辞書には「五山堂詩話 [1807~16]・舎密開宗 [1837~47]・哲学字彙 [1881]・田舎教師 [1909]・歩兵操典 [1992]」の 5 つの年代が記載されたが、最も古いのは「五山堂詩話 [1807~16]」であるため、当該語の出現年代を「1807」と

		している。当該する年代は「10世紀以前」「11世紀」「12世紀」「13世紀」「14世紀」「15世紀」「16世紀」「17世紀」「18世紀」「19世紀」「20世紀以降」の11種があるため、11カテゴリーごとに合致の有無をコード化する。
--	--	---

4. 結果と考察

4.1 RQ1 4種の後続要素の出現頻度

まず、重要漢語動名詞上位50種について、4種の後続要素の出現頻度を調査したところ、図1の結果を得た。

図1 4種の後続要素の比率

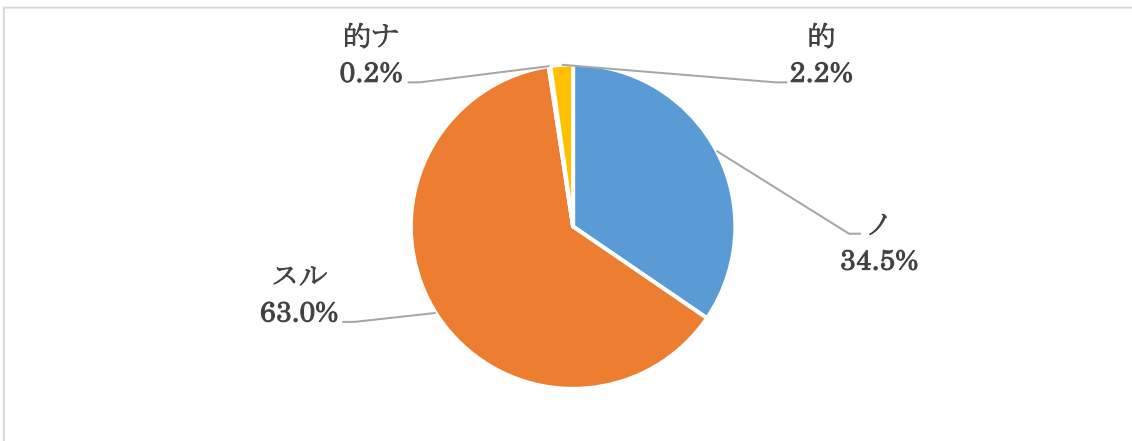


図1からわかるように、4種の後続要素の比率については、「スル」>「ノ」>「的ナ」という順序になっており、「的」と「的ナ」の出現頻度はいずれも1割未満であることが確認された。ここで2点について指摘したい。

まず、1点目は、「スル」は全体の6割以上を占め、圧倒的に多いということである。つまり、他の名詞との結合においては、「漢語VNスルN」が漢語VNの典型的な標準形であると言えるであろう。石(2021)は、動名詞が名詞と動詞の両性を持っているため、論理的には両種の構造が取れるわけだが、「VNスルN」のほうが一般的に成立し、「VNノN」の成立が少ないことが指摘された。今回の結果はこれを支持した。ただ、「ノ」の出現比率も3割に到達し、「スル」に次いで重要な連体修飾の後続要素と言える。

2点目は、「的」と「的ナ」の出現頻度が極めて少ないが、それと結合しうる漢語VNが存在するという点である。鄧(2022)は「安定的(ナ)」は意味的には物事の状態を説明し、形容詞的用法で使われているが指摘された。これによって、「的」あるいは「的ナ」と結合しうる漢語VNは、サ変動詞用法と名詞用法だけではなく、形容詞的用法も持っていると言えるであろう。

以上より、漢語 VN の典型的な連体修飾用法の典型的な標準形は「漢語 VN スル N」であり、「漢語 VN ノ N」も重要な構造であることが確認された。以下、個々の漢語 VN について、4つの後続要素の占有比によって連体修飾タイプを表 2 に示したように分類した。

表 2 各後続要素の出現頻度に基づく 50 種漢語 VN の連体修飾タイプ

タイプ	基準	漢語
「スル」型	「スル」率 60%以上	用意、提供、推進、紹介、実施、注意、使用、設置、確保、対応、存在、作成、実現、検討、規定、活用、理解、期待、入力、説明、購入、参加、発生、確認、展開、利用、工夫、設定、表示、減少 (30 語)
「ノ」型	「ノ」率 60%以上	質問、登録、注目、発表、計算 (5 語)
「的」型	「的」率 60%以上	比較 (1 語)
「的ナ」型	「的ナ」率 60%以上	該当なし
「スル」・「ノ」 両用型	いずれの後続要素は 60%未満が、2つを併 せて 80%を超える	開催、検索、判断、変更、指定、拡大、増加、発見、結婚、表現、選択、開始、評価、変化 (14 語)
「スル」・「的」 両用型		該当なし
「スル」・「的 ナ」両用型		該当なし

表 2 からわかることが 2 点である。1 点目は、上位 50 語の場合、「的ナ」型が存在せず、「的」型が 1 語しかないということである。つまり、漢語 VN について、「的ナ」後続と「的」後続の連体修飾用法が極めて限定的であり、少数の漢語 VN が形容詞的用法を持っているが、その形容詞性は必ずしも高くないと言えるであろう。

2 点目は、上位 50 語のうち、「スル」型が最も多く、「スル」・「ノ」両用型が次に多いということである。これにより、「漢語 VN スル N」の構造は漢語 VN 連体修飾の標準的用法であることを改めて確認した。

以上より、漢語 VN の連体修飾用法の典型的標準形は「漢語 VN スル N」であり、「漢語 VN ノ N」も重要な構造であるが、何らかの条件が存在することが確認された。以下、漢語 VN の持つ各種の言語特性が連体修飾後続要素の選好にどのように影響しているかについて調査する。

4.2 RQ2 「スル」後続に影響を及ぼす要因

すでに述べたように、漢語 VN の持つ言語特性が後続要素の選好に影響している可能性を考

察する。「スル」共起率を目的変数として、Stepwise 変数選択法の重回帰分析を行った結果、用意した 13 種の説明変数のうち、4 つが選ばれ、最終的に次のような回帰式が得られた。

「スル」共起率 = $0.049 + 0.107 * \text{文書ジャンルの出現率} - 0.031 * 18 \text{世紀} - 0.033 \text{VA} - 0.037 * \text{中国語副詞}$ ($p < .001$, $R^2 = .374$)

モデルは有意であり、説明力は 37.4%である。このモデルに従えば、漢語 VN が法律、白書、国会会議録などのような政治・社会に関する公的な文書に出現する場合は、「スル」後続を選択する確率が高い可能性を意味している。つまり、フォーマリティーや専門性が高い文脈において、漢語 VN が「スル」後続を選好することが示唆された。

一方で、出現年代が 18 世紀である場合、あるいは語構成品詞パタンは VA の場合や中国語での品詞は副詞である場合は、他の後続要素の共起率が高くなる可能性がある。

5. まとめ

以上、本研究では、漢語動名詞の連体修飾要素である「スル」・「ノ」・「的」・「的ナ」の許容度を調査した上で、連体修飾後続要素に影響を及ぼす要因を探索した。以下のような結果を得た。

RQ1 (4 種の後続要素の出現頻度) では、「スル」は最も典型的な後続要素であることがわかった。また、後続要素によって「スル」型、「ノ」型、「的」型、「的ナ」型、「スル」・「ノ」両用型の 5 つのタイプに分けられたが、重要漢語動名詞上位 50 語のうち、「的ナ」型が存在していなく、「スル」型が最も多いことが確認された。

RQ2 (「スル」後続に影響を及ぼす要因) では、漢語 VN が法律、白書、国会会議録などのような政治・社会に関する公的な文書に出現する場合は「スル」を選択する確率が高いということが確認された。

ただし、本研究には、(1) 調査対象語の数が限られている、(2) モデルの構築について、考慮した説明変数の種類が限られている、などの課題も残されている。今後の研究では、調査対象語の範囲を拡大し、かつ他の説明変数を考慮しつつ、さらに考察の精緻化を図りたい。

引用文献

- 石立珣 (2021) 「動名詞の連体用法『VN ノ N』の使用実態と成立条件—中国人日本語学習者の使用状況から—」『日本語教育』 179, 77-92.
- 北原保雄 (2007) 『日本国語大辞典 第二版』小学館.
- 桑原洋子 (2017) 「漢字 2 字熟語が漢語動名詞かどうかの判断に及ぼす語構成の影響：非漢字系中上級学習者対象の調査の結果から」『国際教育交流研究』 1, 27-36.
- 国立国語研究所 (2004) 『分類語彙表：増補改訂版』大日本図書.
- 小林英樹 (2004) 『現代日本語の漢語動名詞の研究』ひつじ書房.
- 中国社会科学院語言研究所詞典編輯室 (2016) 『現代漢語詞典』商務印書館.

- 陳迪 (2023) 「書き言葉・話し言葉コーパスデータに基づく高頻度漢語動名詞の品詞性の再考：日本語教育の視点から」『言語資源ワークショップ発表論文集』 1, 215-232.
- 鄧琪 (2022) 「コーパス調査に基づく現代日本語における非和語系語彙の用法解明と日本語教育への応用」神戸大学博士論文.
- 西尾実・岩淵悦太郎・水谷静夫・柏野和佳子・星野和子・丸山直子 (2019) 『岩波国語辞典 第八版 (LogoVista 版)』 岩波書店.
- 松村明・三省堂編修所 (2019) 『大辞林第四版 (LogoVista 版)』 三省堂.
- 山田忠雄・倉持保男・上野善道・山田明雄・井島正博・笹原宏之 (2020) 『新明解国語辞典 第八版 (LogoVista 版)』 三省堂.
- Tony, McEnery & Richard, Xiao (2004). The Lancaster Corpus of Mandarin Chinese. Lancaster University.