

文化審議会国語分科会・言語資源小委員会 発表資料

2025/12/23

## 新聞経年コーパスを用いた外来語の基本語化の研究

金 愛蘭（日本大学文理学部）

## 0. はじめに

発表者は、これまで、20 世紀後半の新聞語彙に「抽象的な意味を表す外来語の基本語化」という現象が生じていることを、大規模な経年コーパスを自ら作成し、それにもとづく計量的な調査と、特定の外来語に注目した事例研究によって実証しようとしてきた（金 2024）。本発表ではその要点をまとめ、言語コーパスの意義や活用法に関する議論に供したい。

## 1. なぜコーパスを作ったのか

1) 外来語の増加は、20 世紀の日本語語彙に生じた最も大きな変化であるといえる。

2) 外来語は基本語彙にも（少しずつ）進出している。

宮島達夫氏は、国立国語研究所の「雑誌九十種の語彙調査」（1956 年）と「月刊雑誌七十誌の語彙調査」（1994 年）の上位 1000 語における語種構成を以下のように集計・比較し、

	和語	漢語	外来語	混種語	合計
90 種	578	386	18	18	1,000
70 誌	499	416	71	14	1,000

外来語の比率が 1.8%から 7.1%に「激増している」こと、すなわち、この 40 年ほどの間に雑誌の基本的な語群において外来語の増加が顕著に進んだことを報告している（宮島 2009）。

3) 具体的な意味を表す外来語だけでなく、抽象的な意味を表す外来語（下線部）も基本語彙に進出（＝基本語化）している。

90 種の外来語（17 語）：ウエスト、カメラ、スカート、スター、スポーツ、ダーツ、チーム、デザイン、テレビ、バス、ファン、ブラウス、ページ、ポケット、ボタン、ホテル、ラジオ、アパート

70 誌の外来語（71 語）：タイプ〔型〕、エンジン、デザイン、レース〔競走〕、モデル、ポイント、テレビ、チーム、ホテル、ページ、スポーツ、システム、セット、プレ

ゼント, クラス, シリーズ, スタイル, サイズ, クラブ〔団体〕, ドア, センター, ジャケット, イメージ, ホール〔コンサート〜〕, メーカー, バランス, コース, バッグ, カラー, スキー, パソコン, ジャパン, データ, アルバム, ボール〔球〕, カップ, パワー, ピアノ, ママ, バス〔車両〕, ボディー, オリジナル, ソフト, オープン, トップ, チェック〔検査〕, コンサート, ゲーム, グループ, シャツ, ナンバー, ジャズ, バイク, テーマ, パンツ, ファン, ビル, イヤリング, ドライバー, プロ〔玄人〕, ライン, メンバー, コーナー, モーター, シンプル, サービス, エネルギー, ラヴ, プログラム, スカート, デビュー

4) 抽象的な外来語の基本語化は、社会や生活の近代化といった言語外的な要因で説明できる具体名詞の場合と違って、言語内的な要因（＝和語や漢語の既存類義語がありながら、なぜ基本語化したのか）を明らかにする必要がある。

5) そのためには、語彙表だけでなく、対象資料の原文を入力したコーパスを経年的に作成して、外来語とその類義語の意味・用法の変化を調査すること（＝計量的語彙史研究）が必要になる。

## 2. どのようなコーパスを作ったのか

1) 大学院在学中に、外来語「トラブル」の用例を 20 世紀後半の『毎日新聞』から 10 年おきに採集することとし、1950～80 年については『縮刷版』を利用した。しかしこれでは、新たな外来語やその類義語を調査しようとするたびに目視による用例採集を行わねばならず、データをコーパス化する必要性を痛感した。

2) 幸い、2008 年に（財）博報児童教育振興会の助成を得て、『毎日新聞』の 1950 年から 2000 年までほぼ 10 年おきに、毎月 2 日分の朝刊全紙面記事の見出しと本文を入力した『毎日新聞経年コーパス』（以下「MKC」）を作成することができた（1950 年・60 年・70 年・80 年は『縮刷版』を、1991 年・2000 年は『CD－毎日新聞データ集』を利用）。

3) その後、このコーパスを 3 回にわたって増補した。2009 年の「第 2 版」では毎月 3 日分に増補した。2014 年の「第 3 版」では 2010 年分を追加し、2018 年の「第 4 版」では、紙面数の少ない 1950 年のみを毎月 9 日分に増やした。

4) 表 1 に、国立国語研究所の形態素解析ツール『茶まめ』を用いて得た最新「第 4 版」のデータ量（短単位・延べ語数）を示す（助詞・助動詞・記号等をすべて含む）。調査各年の平均延べ語数は 193.6 万語となる。なお、国立国語研究所の『太陽コーパス』は 216.4 万語、『昭和・平成書き言葉コーパス』は 345.7 万語である。

表1 MKC（第4版）のデータ量（短単位・延べ語数）（単位：万語）

1950年	1960年	1970年	1980年	1991年	2000年	2010年	計
147.3	142.2	202.5	210.7	208.6	237.0	207.0	1355.3

### 3. コーパスから何がわかったのか

- 1) 新聞においても外来語が基本語彙に進出している。ただし、雑誌に比べるとその速度は遅いようだ。

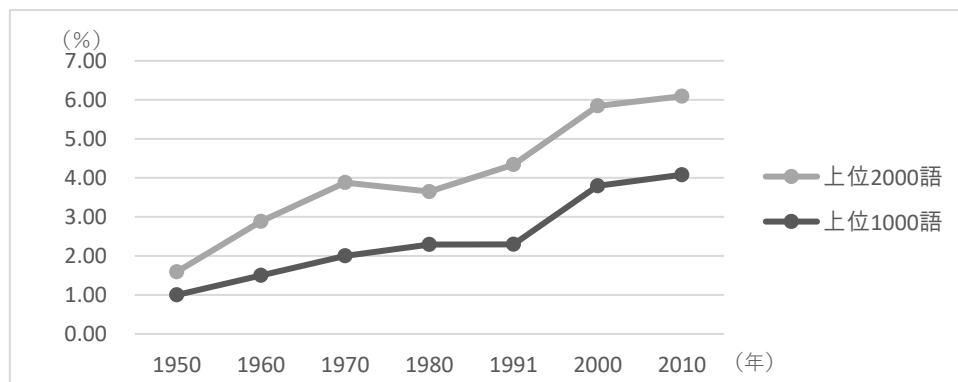


図1 MKC（第4版）上位1000語・2000語中の外来語の割合  
（助詞・助動詞，記号，固有名を除く）

- 2) 基本語彙に進出している外来語には，具体名詞（『分類語彙表』大分類の1.2, 1.4, 1.5）より抽象名詞（同1.1, 1.3）の方が多い。

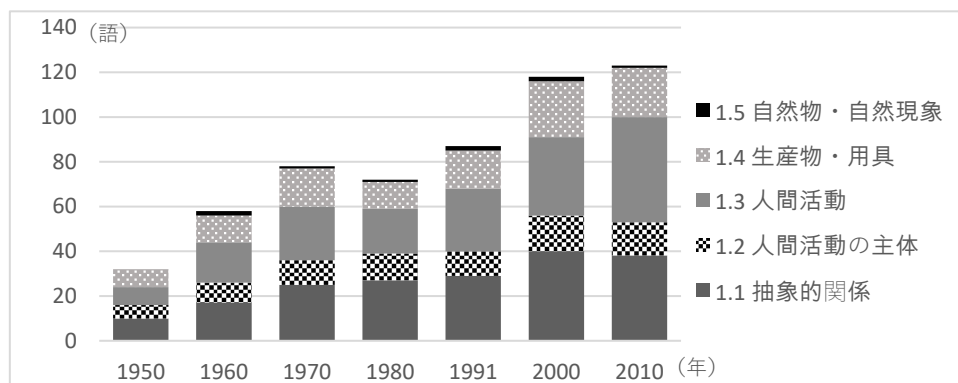


図2 MKC（第4版）上位2000語中の外来語の意味分類（『分類語彙表』の大分類）

- 3) 基本語彙に進出した外来語とその類義語との量的な関係には様々なものがある。図3の「テーマ」は，外来語が優勢な類義語に取って代わるような変化，図4の「ケース」は，外来語が一部の類義語と並び立つような変化を見せている（類義語は，遠藤織枝ほか編『使い方の分かる類語例解辞典 新装版』（2006年，小学館）による）。

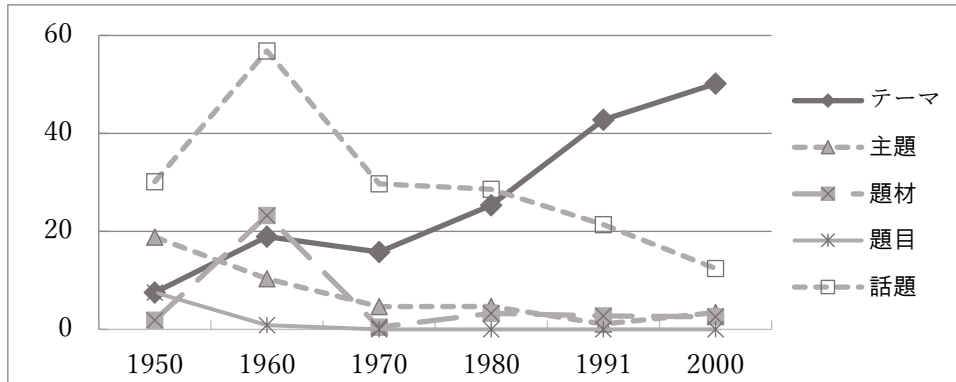


図3 「テーマ」と類義語の量的推移（MKC 第1版，自立語の100万字あたり換算値）

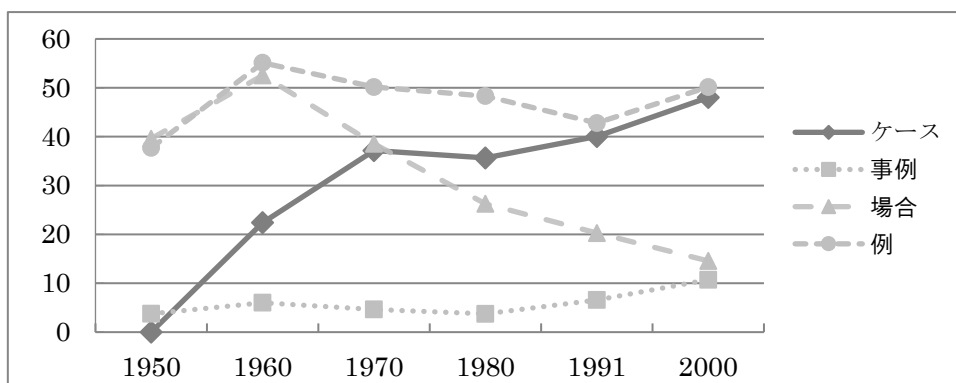


図4 「ケース」と類義語の量的推移（MKC 第1版，自立語の100万字あたり換算値）

4) 多義語化し，類義語の上位語となることで基本語化した外来語がある。

「トラブル」の意味は，「トラブル」がどこに（何に）発生するか，発生した「トラブル」の内容がどのようなものであるかという観点から，以下の3種6類に分類できる。

#### I. ヒトとヒトとのトラブル

- ①〔デキゴトのトラブル〕（例）R容疑者が1週間前にホテルで従業員とトラブルを起こし「火をつける」と騒いだことがあり，……
- ②〔関係のトラブル〕（例）Mさんは交友関係でトラブルを抱えていたとみられることから，捜査本部は交友関係を中心に捜査していた。

#### II. モノのトラブル

- ③〔機械のトラブル〕（例）42秒後，第1段ロケットにトラブルが起き，打ち上げは失敗。
- ④〔身体のトラブル〕（例）春先はにきびなど肌のトラブルが起きやすい季節。

#### III. モノゴトのトラブル

- ⑤〔運営・運用のトラブル〕（例）今月から導入された介護保険で，全国の現場で生じたトラブルについて報告を求めている厚生省は4日，……

- ⑥〔事故・事件のトラブル〕(例) 午前 10 時ごろには新幹線の新大阪駅ホームから乗客が線路内に降りるトラブルがあり、上下 7 本が 10～13 分遅れた。

これらの意味はすべてが初めからそろっていたわけではなく、表 2 のように、1960 年にまずは〔デキゴト～〕が現れ、続いて 70 年に〔関係～〕〔機械～〕〔運営・運用～〕〔事故・事件～〕が、80 年に〔身体～〕が現れており、「トラブル」は 60 年ごろから 80 年ごろにかけてこの順で意味が拡大し、多義語化したことがわかる。

表 2 「トラブル」の意味とその用例数 (MKC 第 0 版, 自立用法)

		60 年	70 年	80 年	91 年	00 年	(類義語)
Ⅰ. ヒトとヒトとの	①デキゴトの	17	43	46	50	41	けんか, 口論, いさかい, 衝突, …
	②関係の		1	13	5	18	もつれ, 不和, 不仲, …
Ⅱ. モノの	③機械の		2	10	4	102	故障, 不調, 不具合, …
	④身体の			4	5	3	不調, 悩み, 疾患, 故障, …
Ⅲ. モノゴトの	⑤運営・運用の		1	3	8	11	障害, 支障, 混乱, …
	⑥事故・事件の		2	12	15	23	事故, 事態, 事件, 不祥事, …
計		17	49	88	87	198	

多義語化によって、「トラブル」はこれら 3 種 6 類に共通する《深刻・決定的な危機的事態に至る可能性を持って顕在化した不正常な事態》という抽象的な意味を表す基本語として成立し、同時にそれぞれの意味の類義語の「上位語」の位置に立つことになった。類義語が分担して表す、より具体的ないし限定的な《事態》は、すべて上位語の「トラブル」で表すことができる。

5) 特定の用法で類義語を大きく上回ることによって基本語化した外来語がある。

図 4 に示すように、「ケース」は「事例」「例」「場合」と類義の関係にある。これら 4 語は、基本的に以下の 4 種の用法をもっている（「場合」は B の用法が少ない）。

- 単独の用法 (例) 男女の賠償額にケースによっては一千万円近い差が生じている。
- 合成語の構成要素 (例) 「特殊ケース」「虐待ケース」「例外的ケース」
- 名詞句の被修飾語 (例) 「仙台地裁のケース」「悪質なケース」
- 連体修飾節構造の被修飾語 (例) 極端な食事制限のストレスから、拒食症や過食症といった摂食障害を引き起こすケースも少なくない。

MKC 第 1 版における「ケース」の用例数の推移を用法別に見ると (表 3), 連体修飾節構

造の被修飾名詞となる用法が大きく増えていることがわかる。

表3 「ケース」各用法の使用頻度（MKC 第1版，実数）

	50年	60年	70年	80年	91年	00年
A. 単独		2	7	6	2	3
B. 合成語の構成要素		4	11	7	8	1
C. 名詞句の被修飾語		10	30	23	19	23
D. 連体修飾節構造の被修飾語		11	33	43	53	91
計		27	81	79	82	118

用法ごとに類義語の「例」「場合」と比較しても（図5－1～4）、「ケース」が連体修飾節構造の被修飾名詞となる用法で類義語を大きく上回るようになってきたことがわかる（MKC 第1版，100万字あたり換算値）。なお、「事例」は用例数が少ないため省略する。

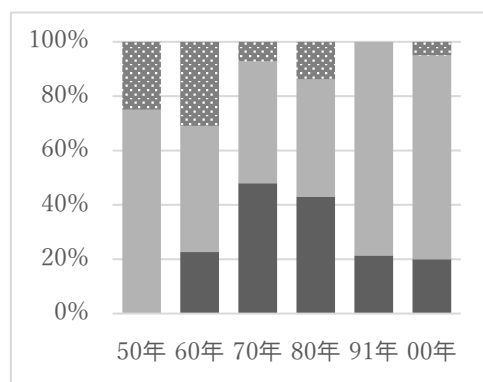


図5－1 A. 単独

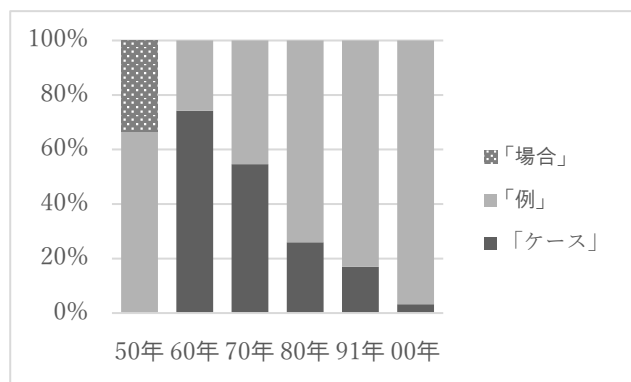


図5－2 B. 合成語の構成要素

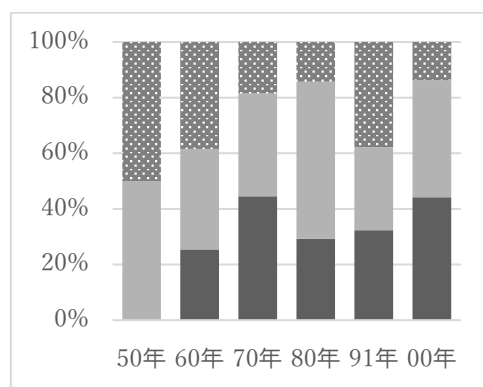


図5－3 C. 名詞句の被修飾語

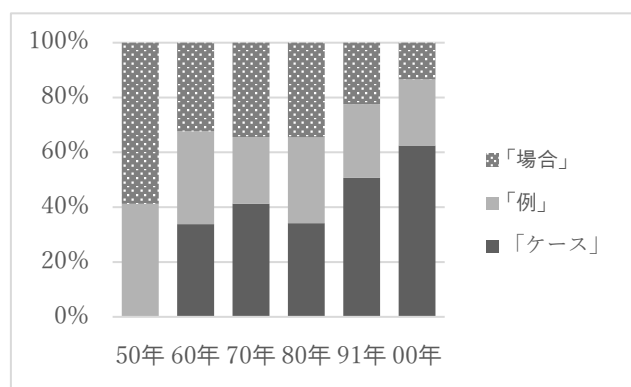


図5－4 D. 連体修飾節構造の被修飾語

「ケース」は、連体修飾節構造の被修飾語（そのほとんどは同格連体名詞）という用法を大きく増やしたことで基本語化したと考えられる。その多くは、下の例のように、「(すでに起こった) 良くないコトガラ (=ケース) が〈多い／有る〉」という表現に用いられている

(225 例中, 112 例)。

(例) だが, 児童相談所が事態を把握しながら救えなかったケースが, 98 年度は判明しただけで 8 件 (8 人) あった。積極的な対応をちゅうちょしたり関係機関との連携がうまくいかないケースが依然多い。

なお, 「事例」は 結合用法 (例: 活用事例), 「例」は助数詞の用法 (例: 300 例) や結合用法 (例: 感染例), 「場合」は提題用法 (例: 今回の場合, ~) や仮定条件的な用法 (例: そのようなことが起きた場合には, ~) において力を発揮しており, 全体として, 形式名詞化する「ケース」との分担を明確化させている。

6) 文章構成機能を担う (テキスト構成語として働く) 外来語が増えている。

“指示詞「この」+名詞”という形式 (以下「コノ語句」) には, 文章内での前方語句の再表現の仕方に, 次のような 3 つの異なる方式が認められる。

①繰り返し: 文章の前方にある語と同じ語をコノ語句でも繰り返して用いる。

(例) 大型人工衛星には予備の燃料を持ったエンジンを装備することもできる。このエンジンによって (略) 燃焼を防ぐことができる。

②言い換え: 文章の前方にある語句をコノ語句で別の語に言い換える。

(例) うつとしい梅雨期, (略) 心身ともにゲンナリする季節だ。(略) しのぎにくいこのシーズンを健康に乗切する方法はないものか。

③捉え直し: 文章の前方で述べられている叙述 (事柄) をコノ語句で名詞化し, 後続の叙述につないでいく。

(例) ローラー・ゲームがテレビで放送されたせいで, 去年は男の子の間にローラー・スケートの売れ行きが好調だった。だが, このブームもすでに峠を越した (略)。

表 4 は, MKC 第 2 版に現われる コノ語句を構成する (単純語の) 外来語 について, これら 3 方式の用例数の推移を集計したものであり, 図 6 は, その構成比の推移を 100 万字あたりの換算値により示したものである。この図をみると, データの規模の小さい 1950 年を別にすれば, 「繰り返し」が減り, 「捉え直し」が増えてきていることがわかる。

表 4 コノ語句による再表現の方式 (実数)

	50 年	60 年	70 年	80 年	91 年	00 年
①繰り返し	9	79	142	125	46	45
②言い換え	2	5	18	16	5	13
③捉え直し	6	12	26	26	17	31

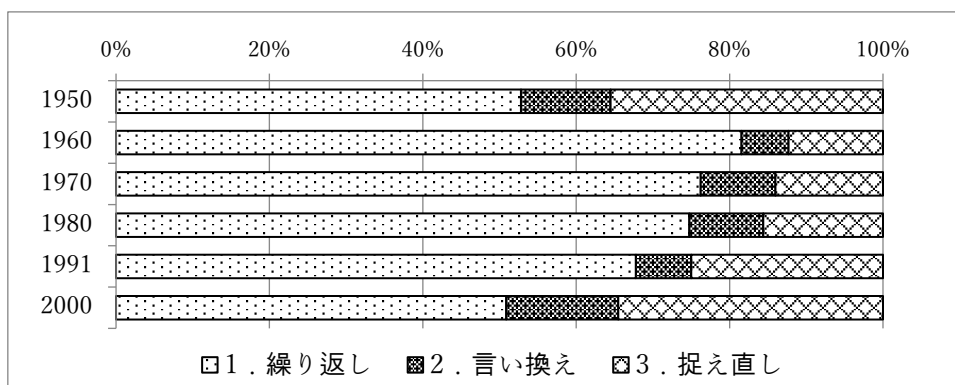


図6 コノ語句による再表現の方式（換算値の比率）

①繰り返しや②言い換えは具体名詞・抽象名詞のどちらでもよいが、③捉え直しは、語より大きな単位で表現される事柄を一語で名詞化するものであるから、基本的には抽象名詞でなければならない。したがって、この結果は、③捉え直しという文章構成機能を担うことのできる——高崎みどり氏の「テキスト構成語」（高崎 2021）として働く——抽象的な外来語が増えていることを示している。

#### 4. なぜコーパスが必要なのか

1) 前節の1)～3)は語彙表のみでも可能だが、4)～6)は原文を入力したコーパスでなければならない。

2) このうち、「トラブル」に見られる「多義語化」や、「ケース」に見られる「既存類義語の一部用法への進出」は、樺島忠夫氏が 2004 年の著作で指摘していることとも重なる。

既存の適切な語によって表現する労力を省いて、覚えた外来語、例えば「アクセス」を乱用すると、その外来語の使用が一般に広がるにつれて、次第に多義語化して、「接続」「利用」「交通手段」「連絡」「参入」のような、文脈上で適切な意味の区別をぼやかしてしまうことになります。(樺島 2004, p.109)

「開く」「けずる」「きる」は、日本語の基本的な語で、これから一〇〇〇年後まで使われ続けるでしょうが、基本的な語の意味の一部を（「カット」という）外来語によって「カット」されることが生じ、この「カット」される用法がさらに大きくなると、日本語にとって基本的な語であっても、ほそぼそと生き残っているに過ぎない状態になり、やがては死語となりかねません。(同, p.118。カッコ内は発表者による注記)

3) 樺島氏によれば、「トラブル」や「ケース」は、「アクセス」や「カット」と同様、「その語を使わなくても、これまでにある日本語で表せる」(同, p.121) 使う必要のない外来

語である。樺島氏は、これら「使う必要のない外来語」に「基本的語彙を侵させるな」(同, p.115)と警告している。

4) しかし、MKC の調査から明らかなように、新聞には、既存の類義語 (=「これまでにある日本語」)があるにもかかわらず基本語化している抽象的な外来語が数多くある。新聞は、なぜ使う必要がない外来語を多用し、それらを基本語化させているのだろうか。

5) 前述したように、「トラブル」は、その多義語化によって《深刻・決定的な危機的事態に至る可能性を持って顕在化した不正常な事態》という抽象的な意味を獲得し、《事態》の詳しい内容がわからず、特定の下位語で表現できない段階でも、とりあえず「(何らかの)トラブル」と書いておけば済ませることができる「便利」な単語となっている。

6) 「ケース」も、連体修飾節構造の被修飾語 (=同格連体名詞) という用法での使用を大きく増やして、「(すでに起こった) 良くないコトガラ (=ケース) が〈多い／有る〉」という、これも新聞で多用される表現に使える「便利」な単語になっている。

7) とすれば、「トラブル」も「ケース」も、新聞にとっては「使う必要のない単語」とは言えないことになる。新聞で基本語化している抽象的な外来語は、新聞の文章構成に「必要」な単語である可能性がある。ただし、そのことを実証するには、適切な経年コーパスを作成し、それにもとづく「計量的語彙史研究」を進めていくことが必要となる。

#### 【文献】

樺島忠夫(2004)『日本語探検－過去から未来へ－』角川書店

金 愛蘭(2024)『外来語の基本語化－現代新聞「叙述語彙」への進出－』大阪大学出版会

国立国語研究所 (1987)『国立国語研究所報告 89 雑誌用語の変遷』秀英出版

高崎みどり(2021)『テキスト語彙論－テキストの中でみることばのふるまいの実際－』ひつじ書房

宮島達夫(2009)「語彙史の比較 (1) 日本語 (雑誌 90 種と 70 誌)」『京都橘大学研究紀要』35, pp. 218-199